Unlocking Practical Cardiac Monitoring Capabilities on True Wireless Stereo Earbuds

Yongjian Fu[1], Ke Sun[2], Xinyu Zhang[3], Ruyao Wang[2], Xinyi Li[4], Hao Pan*[5], Yaoxue Zhang[1], Ju Ren[1,6]

[1]Department of Computer Science and Technology, Tsinghua University, China, [2]University of Michigan, USA, [3]University of California San Diego, USA, [4]Northwest University, China, [5]Shanghai Jiao Tong University, China, [6]State Key Laboratory of Internet Architecture, Tsinghua University, China

Email:{fuyongjian,renju,zhangyx}@tsinghua.edu.cn,{kesuniot,ruyao}@umich.edu xyzhang@ucsd.edu,xinyili@nwu.edu.cn,panh09@sjtu.edu.cn

ABSTRACT

The human ear presents unique physiological and practical advantages for cardiac monitoring. However, existing earable systems rely on extra hardware and complex processing, making them unsuitable for resource-constrained True Wireless Stereo (TWS) earbuds. We present EarCardio, a system that repurposes the built-in IMU sensors of commercial TWS earbuds for cardiac monitoring. Our key insight is that the built-in IMUs can capture ballistocardiogram (BCG) signals, which are streamed to a smartphone via Bluetooth for processing. EarCardio enhances the low-rate and unreliable IMU data through multi-axis fusion and signal enhancement, and further reconstructs fine-grained seismocardiogram (SCG) signals using a region-focused translation framework. We implement EarCardio as a real-time system and evaluate it with 110 human subjects. Results show that EarCardio accurately reconstructs cardiac signals and remains robust to motion artifacts, missing data, and low sampling rates, thus enabling a wide range of cardiac monitoring applications.

CCS CONCEPTS

• **Human-centered computing → Ubiquitous and mobile computing systems and tools**.

KEYWORDS

Cardiac monitoring, BCG, cardiogram reconstruction, earbuds.

*Corresponding author

Yongjian Fu performed this work as a visiting student at the University of California San Diego.

This work is licensed under a Creative Commons Attribution 4.0 International License.

MobiCom '26, October 26–30, 2026, Austin, TX, USA

© 2026 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-2505-0/2026/10

https://doi.org/10.1145/3795866.3796678

ACM Reference Format:

Yongjian Fu[1], Ke Sun[2], Xinyu Zhang[3], Ruyao Wang[2], Xinyi Li[4], Hao Pan*[5], Yaoxue Zhang[1], Ju Ren[1,6]. 2026. Unlocking Practical Cardiac Monitoring Capabilities on True Wireless Stereo Earbuds . In *The 32nd Annual International Conference on Mobile Computing and Networking (MobiCom '26), October 26–30, 2026, Austin, TX, USA.* ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3795866.3796678

1 INTRODUCTION

Cardiovascular diseases (CVDs) are the leading cause of global mortality, accounting for approximately 17.9 million deaths annually [1]. This staggering figure underscores the critical need for innovative solutions in cardiac care, particularly non-invasive, continuous monitoring technologies that can enhance early detection of CVDs while alleviating the burden on healthcare systems. The widespread adoption of TWS earbuds, expected to reach over 700 million users by 2026 [2], offers a promising mobile health platform to meet this demand. The ears' anatomical structure, rich in vital blood vessels like the superficial temporal and posterior auricular arteries, provides a stable site for accurate cardiac monitoring.

Current research in earable computing leverages these vascular advantages to develop cardiac monitoring solutions [3–6]. Representative techniques include integrating photoplethysmography (PPG) sensors for in-ear PPG monitoring [3] and using ultrasonic transducers to measure cardiac dynamics [5]. Headphone drivers have also been repurposed [4, 6] to detect heartbeats and reconstruct phonocardiogram (PCG) signals. However, these solutions necessitate additional hardware design or complex software processing. This poses significant challenges for commercial TWS earbuds, which are constrained by their form factor and resources, including energy, communication, and computation capabilities.

In this paper, we explore the potential of *repurposing commercial TWS earbuds for cardiac monitoring applications.* Our preliminary studies demonstrate that standard IMU sensors, *i.e.*, accelerometers and gyroscopes, embedded in TWS

earbuds and originally intended for head motion tracking, can potentially capture in-ear BCG signals. Based on this observation, we propose EarCardio, a novel system that augments TWS earbuds with BCG monitoring capabilities. EarCardio can convert the acquired BCG signals into SCG waveforms that are associated with clinically relevant cardiac metrics. It is compatible with COTS TWS earbuds, *e.g.* Apple AirPods [7] and eSense [8], and can be implemented through software enhancements, without any hardware modifications.

To achieve these advantages, we address three challenges:

Challenges of reusing the limited TWS earbuds hardware and software. To preserve the earbuds' normal functions and minimize software overhead, EarCardio utilizes the existing Bluetooth Low Energy (BLE) communication channel to stream raw IMU data to a mobile device (*e.g.* a smartphone), which then processes the IMU data to derive BCG signals. However, the BLE-based IMU stream is primarily designed for real-time head motion tracking, which tends to be deprioritized by the TWS earbuds and allocated only minimal resources. Thus, the IMU data stream suffers from low sampling rate and unreliable sampling issues, especially when contending with more audio streaming functions. To overcome this challenge, we use the timestamps in the IMU stream to identify the locations of missing samples. We then develop a neural-based cardiographic continuity enhancement method by employing data fusion across multiple axes and ensemble learning from consecutive beats to fill in the missing data. Additionally, we design a super-resolution scheme to reconstruct cardiographic signals at a high sampling rate.

Challenges in reliable in-ear BCG sensing. The BCG signals extracted from IMU typically exhibit low SNR compared to those from specialized sensors, especially in the presence of motion artifacts. Moreover, variability in ear shapes and vessel locations results in different sensitivities across the axes. These in-ear IMU characteristics imply that prior denoising designs developed for more stable signals may not directly transfer to commercial TWS earbuds. To address these challenges, we introduce a two-stage denoising solution. We first mitigate the motion artifacts in the frequency domain using stationary wavelet transform (SWT). Then we design a multi-head attenuation network to select and fuse the axes that are more sensitive to BCG signals. In addition, to enhance model generalization, we leverage an IMU simulator to create highly diverse training data, incorporating different motion artifacts and user profiles.

Challenges of utilizing in-ear BCG for practical applications. The ultimate goal of EarCardio is to harness in-ear BCG for practical applications, particularly in CVD monitoring. However, existing neural-based signal enhancement solutions inadvertently suppress anomalous heartbeat signals. Similarly, anomalous heartbeat peak magnitude might be mistakenly normalized due to biases inherent in the training dataset. To maintain the fidelity of crucial cardiac features, we carefully design our neural networks to achieve super-resolution BCG waveform reconstruction without changing the interval of the signals, thereby preserving essential time-domain information. Additionally, we develop a cardiac-focused loss function that prioritizes the accurate reconstruction of peak regions of the cardiographic signals, while diminishing focus on less critical areas. To further enhance accuracy, we design a transformer-based neural network that fuses multi-axis in-ear BCG signals and reconstructs physiologically aligned SCG [9]. This is motivated by the fact that both BCG and SCG capture cardiomechanical events within the same cardiac cycle, sharing waveform structures rooted in mechanical heart motion, which makes SCG a natural target for translation.

We have implemented EarCardio end-to-end as Android and iOS apps[1], running on smartphones which acquire and process the IMU data from the TWS earbuds. We further collected a real-world dataset consisting of 40 hours of in-ear BCG recordings from over 110 participants, and synthesized an additional 150 hours of data with 202 types of motion artifacts for model training. Our evaluation results show that EarCardio effectively reconstruct SCG signals, with an average cosine similarity of 0.92 across participants with varying BMI, ages, and genders. EarCardio estimates heart rate (HR) and inter-beat interval (IBI) with average errors of 2.73% and 1.74%, respectively, even in the presence of significant motion artifacts, missing data, and low sampling rates. EarCardio operates in real-time on COTS TWS earbuds and paired smartphones with power consumption of only 1.14 mAh/h and 7.26 mAh/h, respectively. Our case studies reveal that the recovered SCG signals can support various healthcare and HCI applications, including long-term HR and Heart Rate Variability (HRV) monitoring, biometric user authentication, pathological case detection, Blood Pressure (BP) estimation, and electrocardiogram (ECG) reconstruction.

We summarize our contributions as follows:

- We conduct comprehensive preliminary studies and demonstrate that standard IMU sensors embedded in TWS earbuds can effectively capture in-ear BCG signals.
- We develop an end-to-end, software-only BCG sensing scheme that can be seamlessly integrated into COTS TWS earbuds without requiring any hardware modifications while minimizing software overhead.
- We design a neural enhancement network to mitigate the motion artifacts and enhance cardiographic signal continuity. We further devise a region-focused signal reconstruction method that fuses and translates multi-axis, coarse-grained in-ear BCG signals into fine-grained SCG signals.

[1]The implementation is available at https://github.com/fuyongjian/EarCardio.

Cardiac Signals	Measurement Methods	Frequency Range	Sensing Modality	Sensor
PPG	Blood volume	0.5~4 Hz	Light	LED + PD
ECG	Electrical activity	0.5~100 Hz	Electrical	Electrode
PCG	Heart sound	20~2k Hz	Acoustic	Mic
SCG	Chest vibration	1~50 Hz	Mechanical	Acc/Gyro
BCG	*Body recoil due to ejection of blood*	*1~20 Hz*	*Mechanical*	Acc/Gyro

Table 1: Sensing mechanisms for cardiographic signals.

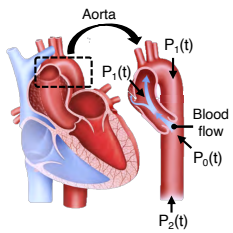


Figure 1: The causes and model of BCG formation.

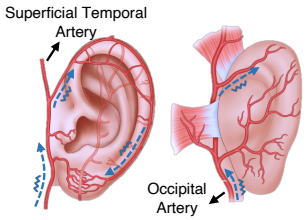


Figure 2: BCG propagation near the ears.

• We conduct comprehensive experiments to demonstrate EarCardio's accuracy, reliability, and effectiveness in supporting diverse cardiac use cases.

2 BACKGROUND AND RELATED WORK

2.1 A Primer on Cardiographic Signals

Cardiographic signals are categorized into four types based on sensing modalities, as summarized in Table 1. PPG utilizes a light source and a photodetector to measure blood volume changes. ECG captures the heart's electrical activity using electrodes. PCG employs microphones or specialized sensors to record heart sounds. SCG detects subtle chest wall vibrations from the heart's mechanical actions, particularly the movement of the heart muscle and the ejection of blood. The BCG used in EarCardio is akin to SCG, measuring the body's recoil movements caused by the heart's mechanical activity during blood ejection, as shown in Figure 1. Below we compare these cardiographic signals from four perspectives.

Clinical Advantages of Extended Monitoring: Routine cardiac check-ups (10–30 s) often miss transient abnormalities such as arrhythmias or subtle episodes of heart failure. Clinical studies show that long-duration or repeated monitoring sessions (from minutes to hours each day) substantially improve event detection compared to isolated spot checks [10, 11]. Cardiac monitoring over extended periods also helps identify mechanical dysfunctions in heart failure [12, 13]. *EarCardio enables practical in-ear measurements that offer opportunistic multi-hour recordings during natural earbud use. This can provide richer coverage than brief check-ups and capture transient events more reliably in a non-intrusive manner.*

Invasiveness and Accessibility: ECG, SCG, and PCG are considered relatively invasive. ECG requires electrodes to be firmly attached to the skin, limiting their use for continuous cardiac monitoring [14]. SCG and PCG generally require attaching sensors to the subject's chest. In contrast, PPG and BCG are less invasive and more accessible through wearable sensing. PPG sensor is a standard module on smartwatches [14], and BCG can be measured using ubiquitous accelerometers positioned on the body or an adjacent surface, such as a bed or chair [15]. *This minimally invasive nature and the reuse of existing sensors inspired the design of EarCardio.*

Medical Applicability and Signal Translation: ECG and PCG are considered clinical gold standards for CVD diagnostics due to the tight sensor attachment and hence highly sensitive measurement. Recent studies highlight that the less-invasive PPG and SCG technologies also closely correlate with myocardial activities [16]. These technologies offer potential for translating PPG/SCG to ECG [9, 17–22]. Despite the potential, they largely depend on controlled, high-quality datasets, which may limit practical applicability [9, 19, 23, 24]. *In contrast, EarCardio captures in-ear BCG signals characterized by low SNR, low sampling rates, unreliable sampling, and significant motion interference. Our case studies involving 100 participants demonstrate EarCardio's effectiveness in real-world cardiac monitoring applications.*

2.2 Earable Cardiographic Signal Sensing

Table 2 compares existing earable cardiographic sensing solutions. In-ear PPG earbuds integrate LEDs and photodiodes to track blood-volume changes for respiration and blood-pressure estimation [25]. While in-ear PPG in TWS earbuds is becoming increasingly feasible (*e.g.*, AirPods Pro 3 [26]), PPG remains an optical hemodynamic modality: it primarily reflects low-frequency blood-volume pulsations and provides limited access to cardiomechanical timing and structure. More critically, PPG does not directly capture cardiomechanical events (*e.g.*, valve-motion–related signatures and systolic timing landmarks) that support systolic-interval/SCG timing and BP-related metrics; these mechanical landmarks can also serve as anchors for feasibility ECG-like reconstruction, motivating complementary modalities rather than replacing PPG.

HeadFi and Asclepius repurpose headphone drivers as PCG sensors [4, 6], but their wired prototypes and auxiliary analog circuitry limit translation to ultra-compact TWS earbuds. Prior work on TWS earbuds instead leverages COTS in-ear microphones or speaker–microphone pairs for sensing, including in-ear PCG [27, 28] and APG [5]. However, in-ear PCG on commodity ANC earbuds is often impaired by microphone-chain filtering and music interference [5], while APG relies on ultrasonic probing via ANC speakers that remains detectable by feedback microphones during playback. Despite this robustness, APG produces PPG-like signals dominated by vascular pulse dynamics (HR/HRV). In contrast, EarCardio offers a software-only approach using only built-in IMUs

	Cardiac Signal	Sensor	HW Change	Monitoring Functionality
In-ear PPG [3]	In-ear PPG	LED + PD	Yes	HR, HRV, BP
HeadFi [4]	Heart Rate	Speaker driver	Yes	HR
Asclepius [6]	In-ear PCG	Speaker driver	Yes	PCG for auscultation
In-ear PCG [29–31]	In-ear PCG	In-ear microphone	No	HR, HRV
APG [5]	In-ear echo	In-ear speaker & microphone	No	HR, HRV, IBI
Around-ear BCG [32, 33]	BCG	IMU sensor	Yes	HR, HRV
EarCardio	**In-ear BCG**	**IMU sensors**	**No**	**HR, HRV, IBI, SCG ECG, and BP**

Table 2: Comparison between various in-ear cardiographic sensing techniques (HW: Hardware).

to extract in-ear BCG and reconstruct SCG/ECG, capturing cardiomechanical features (*e.g.*, valve-related events and systolic intervals) that are largely orthogonal to acoustic cues and complementary for future multimodal fusion.

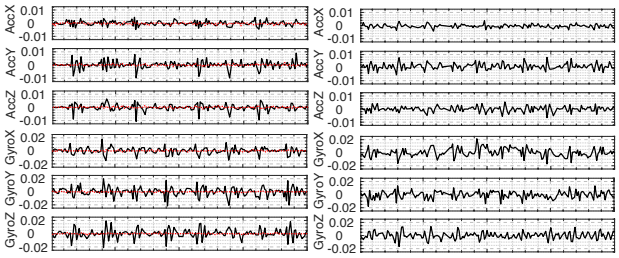
Prior ear-area BCG systems either rely on customized, high-rate IMUs or treat BCG as a complement to in-ear PPG, thus largely sidestepping COTS TWS constraints. FaceReader mounts ~1kHz IMUs on AR/VR headsets[32] but, due to the sensor–artery standoff, the signal is dominated by skull-propagated micro-vibrations and effectively low-pass filtered [34], suppressing valve-motion and systolic timing cues and limiting it to HR estimation. A custom hearing-aid design places a > 500Hz IMU behind the ear[33] with higher fidelity, but requires a nonstandard form factor, added circuitry, and wired links. Recent systems fuse ~500Hz BCG with in-ear PPG for multimodal monitoring[35, 36]. In contrast, COTS TWS earbuds must operate with ~25 Hz IMU sampling, strong motion artifacts, and low-bandwidth/unreliable transmission, while supporting richer analytics such as beat-to-beat SCG/ECG reconstruction and downstream tasks (*e.g.*, authentication and BP estimation).

To our knowledge, EarCardio is the first system to use the built-in, low-rate IMU of COTS TWS earbuds to reconstruct in-ear BCG and SCG/ECG waveforms without hardware changes, protocol tweaks, or noticeable system overhead.

3 PRELIMINARY STUDY

Figure 1 shows the basic in-ear BCG model, which can be interpreted as the blood-pressure gradients in the ascending and descending aorta [37]. When an instantaneous force is exerted on the blood in the main artery, the BCG force F_{BCG} can be estimated as

$$F_{BCG}(t) = S_d[P_1(t) - P_2(t)] - S_a[P_0(t) - P_1(t)] \quad (1)$$



(a) Left-ear BCG from #A (b) Right-ear BCG from #B

Figure 3: Raw IMU waveform collected by Apple AirPods Pro (Duration: 8 seconds; Sampling Rate: 25 Hz; Units: Acc m/s^2 and Gyro dps). Red dotted line represents the noise floor when the earbud is placed on a flat surface.

where S_d and S_a are the average cross-sectional areas of the descending and ascending aorta, respectively. $P_0(t)$, $P_1(t)$ and $P_2(t)$ represent the blood pressure waves of the ascending aorta inlet, arch outlet/inlet, and descending aorta outlet, respectively. As shown in Figure 2, two arteries around the ear, *i.e.* superficial temporal artery and occipital artery, introduce significant mechanical vibration inside the ear. *Our key finding is that in-ear BCG signals can propagate into the ear canal, allowing them to be measured by the IMU sensors in COTS TWS earbuds.*

3.1 Feasibility of In-ear BCG

To assess the feasibility of in-ear BCG, we conducted preliminary studies using two commercial TWS earbuds that provide access to raw accelerometer and gyroscope data from their IMU sensors, *i.e.* the Apple AirPods series [7] and the eSense open earable platform [8]. During the experiment, subjects were instructed to wear the earbuds as they normally would while remaining stationary. Figure 3 shows the raw IMU waveform recorded by Apple AirPods Pro (2nd generation), which exhibits the same pattern to typical BCG signals [37]. We repeated this experiment using eSense and observed the same. This attributes to the high sensitivity of IMU sensors. According to the datasheet, earable accelerometers have a typical measurement range of $\pm 2g$ and a digital resolution of 16 bits, translating to a resolution of $0.06mg/LSB$. The gyroscopes feature a measurement range of $\pm 125dps$ (degrees per second) and a digital resolution of 16 bits, resulting in a resolution of $0.004dps/LSB$. Our preliminary study indicates that the average maximum peak amplitude of in-ear BCG, characterized by accelerometer and gyroscope, is approximately $0.005m/s^2$ and $0.03dps$, respectively, resulting in an average empirical peak SNR of $38dB$ and $19dB$. *This suggests that the IMU sensors in TWS earbuds have sufficient resolution to sense in-ear BCG signals.*

To measure the reliability of in-ear BCG, we employ the frequency domain Bandpass SNR [38], which evaluates how effectively the $1 \sim 10Hz$ cardiac frequency components are

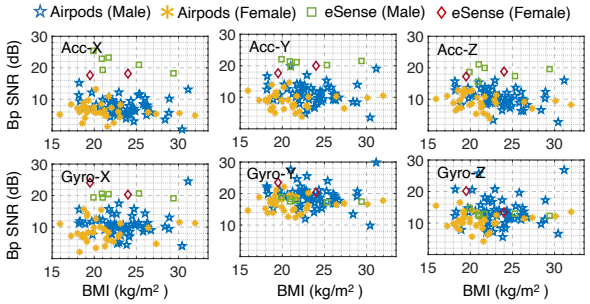


Figure 4: Frequency domain bandpass SNR benchmark across different IMU axes and subjects.

preserved amidst noise. The Bandpass SNR (in dB) is calculated as: $10\log_{10}(P_{signal}/P_{noise})$, where we compute the power spectral density (PSD) to determine the mean power of both the signal (P_{signal}) and noise (P_{noise}). P_{noise} is sampled when placing the earbuds on a stationary surface.

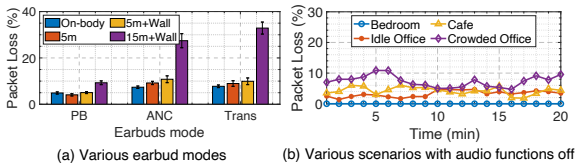
We measured the SNR of multi-axis in-ear BCG signals across 100 (Figure 4). Section 5 elaborates on our participants' profiles. Our findings are and summarized as follows.

Variable SNR across different IMU axes. We observe significant variation in SNR across different IMU axes. The axis oriented toward the occipital artery (Y-axis for Apple AirPods and X-axis for eSense) achieves the highest SNR, averaging $18.21dB$ and $20.51dB$, respectively. Conversely, the axis facing outward from the subject (X-axis for Apple AirPods and Z-axis for eSense) exhibits the lowest SNRs, with averages of $10.04dB$ and $13.93dB$, respectively.

Consistent SNR across subjects of different gender, BMI, and age. Female subjects exhibit slightly lower SNRs than male subjects, with an average difference of $1.53dB$ across different axes. BMI affects the SNR, but insignificantly, with only a $1.87dB$ reduction observed between underweight (BMI $< 18.5kg/m^2$) and obese subjects (BMI $> 30kg/m^2$). With respect to age, the average SNR is $12.33dB$ for subjects under 30 years of age, $10.67dB$ for those between 30 and 40 years, and $9.56dB$ for those over 40 years. This decrease in SNR with age is likely due to the higher BMI typically associated with older subjects.

Consistent SNR across different ear tips and wearing styles. The SNR is consistent for different ear tip size and wearing styles. Apple provides four ear tip sizes and a fit test to select the best size for users [39]. When users chose an unsuitable ear tip without passing the ear tip fit test, the SNR drops by only $1.34dB$ and $1.13dB$ for the accelerometer and gyroscope, respectively.

Consistent SNR across different COTS TWS earbuds. The IMU sensors in AirPods are not configurable, whereas eSense provides IMU sensor configuration. We configure the eSense to maximize SNR, by minimizing the measurement range of accelerometer with $\pm 2g$ and gyroscope with $\pm 250dps$, respectively, while disabling the on-chip low-pass filter. eSense achieves an average SNR of $18.84dB$ compared to Airpods



(a) Various earbud modes (b) Various scenarios with audio functions off

Figure 5: IMU stream BLE channel packet loss rate.

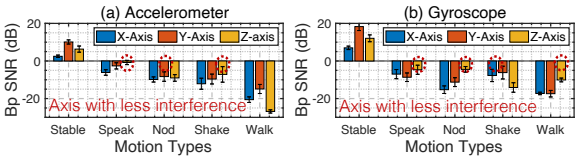


Figure 6: In-ear BCG SNR under common motions.

Pro's $13.12dB$. We suspect that this is also because eSense can be better secured on the subject's ears as it offers specific designs to ensure a firm attachment [40].

3.2 Challenges of In-ear BCG

Despite the potential of in-ear BCG sensing, several key challenges must be addressed for practical usage.

3.2.1 Inherent Challenges due to Earbuds' Hardware and Software Constraints. Because of the limited computational resources, the earbuds' IMU data have to be streamed to an associated mobile device for post processing, which involves two challenges.

Low Sampling Rate. Standard IMU sensors typically support sampling rates exceeding $500Hz$. However, TWS earbuds are limited to much lower rates due to BLE constraints. COTS TWS earbuds limit each BLE packet to transfer one 6-axis IMU data sample without buffering to guarantee real-time processing. The minimal packet interval for Bluetooth 5.4 is $7.5ms$, theoretically allowing a maximum rate of $133.3Hz$. In practice, factors such as energy consumption and computational limitations further reduce this rate. For example, Apple AirPods Series is configured with a minimal packet interval of $40ms$, resulting in a $20 \sim 25Hz$ rate. These rates are sufficient for head motion tracking, but inadequate for cardiac monitoring and analysis, which typically demands a minimum sampling rate of $100 \sim 200Hz$ [41].

Unreliable Sampling. The IMU stream transmitted over BLE exhibits inconsistent sampling intervals. Using AirPods Pro paired with an iPhone (Figure 5(a)), we observe timestamp gaps that fluctuate by multiples of 40ms (nominally 25Hz), indicating missing samples. We identify three root causes. First, resource contention with real-time audio functions degrades sampling reliability, and the loss is most pronounced in ANC and Trans modes. Second, BLE channel conditions matter: when the phone is within ~5m of the user or carried on-body, loss is minimal, whereas at ~15m with wall obstructions, packet loss exceeds 30%. Third, even in close proximity with audio disabled, environmental interference (e.g., nearby BLE devices) can destabilize the link and trigger loss, as shown in Figure 5(b). Finally, the earbuds' low-latency

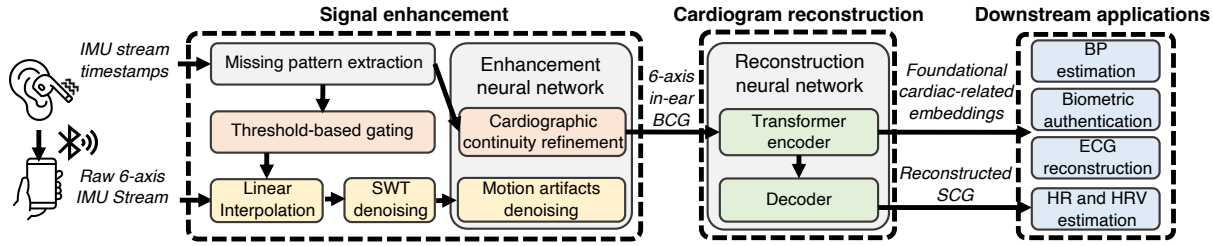


Figure 7: System overview.

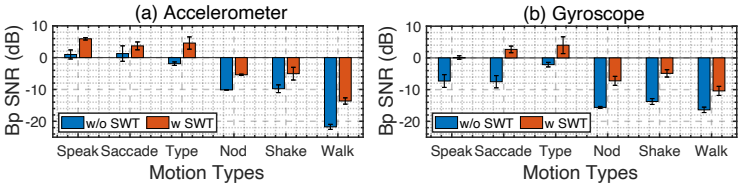


Figure 8: SNR w/ and w/o SWT denoising.

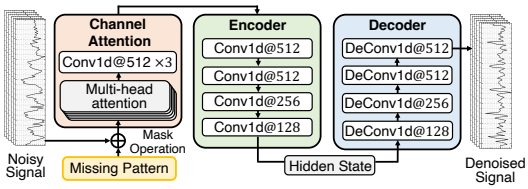


Figure 9: Deep denoising model architecture.

head-tracking and spatial-audio control requires a short BLE supervision timeout (100ms), which further exacerbates loss under congested wireless conditions and aligns with prior BLE wearable measurements[42].

3.2.2 *Challenges in Practical Usage Scenarios.*

Low SNR. We compared the SNR of in-ear BCG with that of other body parts. During the measurement, users were instructed to firmly attach an AirPod earbud to the "Back of the ear", "Chest", and "Neck" using tape. In-ear BCG has $7.97dB$, $6.76dB$ and $2.27dB$ lower SNR than "Back of the ear", "Chest", and "Neck", respectively. This is because the two near-ear arteries are not proximate to the in-ear measurement site (Figure 2). Besides, in-ear BCG were captured with normally worn earbuds, rather than the tape-secured setting.

Motion artifacts. We further quantified the SNR of in-ear BCG across several common motions, including "Speak", "Nod", "Head shake", and "Walk", as illustrated in Figure 6. "Speak" motion introduces interference at intensities comparable to that of stable in-ear BCG. In contrast, head movements and walking produce substantial motion noise, surpassing the desired BCG signals by several orders of magnitude.

Variation across different subjects. Although different subjects achieve similar SNR, their in-ear BCG waveforms exhibit significant variability. Figure 3 demonstrates an example for two subjects. Subject #A's in-ear BCG is sensitive to the Y-axis and Z-axis of the accelerometer and the Y-axis and Z-axis of the gyroscope, while Subject #B is sensitive only to the X-axis and Y-axis of the gyroscope. This variability is attributed to differences in ear shapes, vessel locations, and wearing styles among subjects, which affect the waveform and the sensitivity of different IMU axes.

4 SYSTEM DESIGN

In this section, we introduce the key design components of EarCardio which follow the workflow in Figure 7.

4.1 Motion Artifacts Denoising

Although motion artifacts obscure the subtle in-ear BCG in the time-frequency domain, as visualized in Figure 10(b), two key factors offer new opportunities to denoise such interference. *First*, in-ear BCG has a unique cardiac frequency pattern distinct from arbitrary body motion. This pattern enables us to determine whether the IMU signals are dominated by motion artifacts or in-ear BCG, allowing for targeted enhancement of the relevant signals. *Second*, certain sensors and axes among the 6-axis IMU experience less motion interference during specific motions, as highlighted in the example in Figure 6. Our motion artifact denoising pipeline leverages these two findings, comprising the following two stages.

4.1.1 *SWT-based denoising.*

Common motion artifacts, particularly those generated by slow head movements and systemic body motions, typically have frequency components below $2\,Hz$ [43]. In contrast, the critical peaks of BCG signals, including the *H, I, J, K* and *L* peaks, occur at frequencies above $3Hz$ [44]. Our SWT-based algorithm is designed to reduce the low-frequency dominant motion interference while maintaining the critical high-frequency BCG signals.

The SWT-based denoising method can be expressed as

$$s(t) = \sum_{i=1}^{J} D_i^b(t) + \sum_{i=J+1}^{N} D_i^m(t) \tag{2}$$

where D_i^b and $D_i^m(t)$ represent the ith wavelet coefficients corresponding to high-frequency BCG signals and lower-frequency motion artifacts, respectively. N is the order of SWT, and J is the decomposition level separating the frequencies of motion artifacts and BCG signals. We set $N = 5$ and $J = 2$ when the sampling rate of IMU sensor is $f_s = 25$ Hz. Therefore, the filtered signal frequency band is $> f_s/2^{J+1} = 3.25$ Hz, ensuring the preservation of critical peak information in BCG signals. In practice, we first decompose the IMU signals using SWT, and then reconstruct the enhanced time-domain cardiac signals $s_e(t)$ by applying inverse SWT to

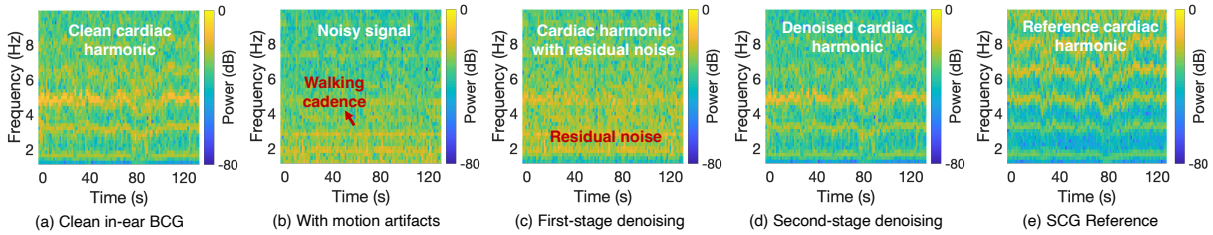


Figure 10: Time-frequency domain visualization for motion artifacts denoising.

high-frequency component $\{D_1^b, D_2^b\}$. Note that these steps are performed separately for each axis.

The key advantage of SWT compared to traditional discrete wavelet transform is its ability to decompose signals into different frequency bands without downsampling. This ensures that the processed signals retain the time-domain details of the BCG signals. However, due to the aforementioned problem of unreliable sampling, directly applying SWT to the raw IMU data would significantly corrupt both the time and frequency information. Fortunately, the IMU stream records the timestamp of each received IMU sample, allowing us to identify the missing patterns in the data stream. Thus, before applying SWT denoising, we first perform a simple linear interpolation, which upsamples the IMU stream to a *stable sampling rate* by filling in the missing data. Note that this step alone is insufficient to fully address the unreliable sampling. It has to be augmented by the neural-enhanced cardiographic continuity algorithms (Section 4.2).

Figure 8 shows the SNR with and without SWT-based denoising under various motion artifacts. The denoising effectively reduces low-frequency motion interference, notably during subtle, slow motions like "Speak", "Saccade" and "Type". However, significant residual noise remains in axes with high-frequency interference, mostly due to fast motions like "Nod", "Shake", and "Walk" (see Figure 10(c)). We note that other potential vibration sources (*e.g.*, earbud speaker playback) are also dominated by higher-frequency components and have limited overlap with the low-frequency cardiac band captured by IMU, and thus are less impactful in practice.

4.1.2 Neural-based multi-axis denoising. We further design a model to reduce the residual noise. It takes the SWT-denoised IMU data as input, and fuses the multi-axis data.

Model design. Figure 9 illustrates our model architecture. It takes 5-second, 6-axis IMU sensor data as input and outputs the corresponding denoised 5-second, 6-axis in-ear BCG signals. To preserve critical timing information, both input and output signals are maintained at 25 Hz to match the IMU sampling rate. Our model consists of two components. *First*, We design a channel-attention module that uses multi-head attention to estimate time-varying axis reliability, selectively up-weighting high-SNR axes/time segments to denoise low-SNR axes via cross-axis evidence aggregation. *Second*, we incorporate a denoising autoencoder, which encodes the input signal into a lower-dimensional representation, effectively filtering out noise while preserving essential features. A decoder then reconstructs the signal, focusing on the cardiac components while suppressing residual motion artifacts.

Training data synthesis. Training a robust denoising model requires large-scale *paired* in-ear BCG data with and without motion artifacts, which is difficult to acquire in practice, even with additional sensors on the earbuds. We address this challenge by observing that motion artifacts act as largely independent and additive interference to in-ear BCG, with minimal impact on blood pulsation [45]. Accordingly, we first collect clean in-ear BCG signals under stationary conditions, separately simulate motion-induced IMU signals, and then synthesize paired training data via linear combination.

To collect clean in-ear BCG signals, we reuse the 100 subjects and instruct them to remain seated and stationary, yielding over 145,000 beats (25 hours). To model motion artifacts, we synthesize motion-induced earbud IMU signals using an IMU simulator, avoiding the need for large-scale real motion data collection. Specifically, we extend the simulator in [46] and drive it with AMASS [47], which aggregates 15 motion-capture datasets and provides 63 hours of high-resolution SMPL meshes covering diverse daily activities. We downsample AMASS to 25 FPS to match TWS sampling, place a virtual IMU at the ear vertex of the SMPL mesh, and generate earbud-frame orientations and accelerations. Training inputs are synthesized by linearly combining simulated motion signals with clean in-ear BCG, resulting in a 150-hour dataset spanning 300+ body shapes and diverse motions, which enables training a generalizable denoising network.

4.2 Cardiographic Continuity Refinement

Our signal enhancement method aims to further address the challenges of cardiographic discontinuity caused by unreliable sampling and data loss. Figure 11 shows a real-world example of an IMU stream received by a smartphone when a stationary subject wore the earbuds in ANC mode. Even after applying linear interpolation, the in-ear BCG signals remain significantly disrupted, particularly when critical peak information is lost, as highlighted in Figure 11(b).

Model design. EarCardio's cardiographic continuity refinement solution builds on two observations. *First*, the peak

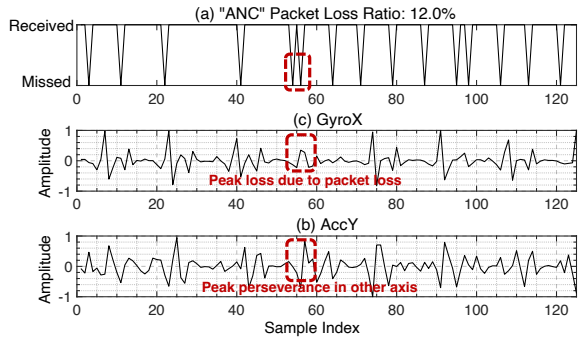


Figure 11: Example of missing data issue.

positions in BCG signals vary across different axes due to the transmission of vibrations through blood vessels in multiple directions [48]. Therefore, even if a specific axis (Figure 11(b)) misses critical peak information, other axes can preserve their own (Figure 11(c)). *Second*, consecutive beats exhibit similar BCG signals and are unlikely to miss the same peak information simultaneously. Thus, they can be combined to resolve cardiographic discontinuities. These two principles align closely with the neural-based multi-axis denoising model. Therefore, the same neural network in Section 4.1.2 can be repurposed to achieve dual objectives: motion artifact denoising and cardiographic continuity refinement. To guide refinement, we derive a missing-sample pattern from timestamp gaps Δt by first computing $k = \Delta t * f_s$, where f_s is the sampling rate, $k = 1$ indicates no loss and $k > 1$ indicates $k-1$ missing samples. We incorporate this pattern also as the input of the model. As shown in Figure 9, EarCardio employs the missing pattern as a binary mask, providing supplementary positional information for multi-head attention and guiding the autoencoder to recover the missing data.

Training data synthesis. To train this model, we need to collect in-ear BCG signals from TWS earbuds without data loss, which serves as the ground truth. To achieve this, we disable all audio functionalities and ensure that no additional BLE or Bluetooth devices are detected nearby. Data is collected while the associated smartphone receiver is placed on a table nearby, as shown in Figure 13. After data collection, we manually remove any received IMU streams with packet loss, creating a clean dataset at a 25 Hz sampling rate. We simulate packet loss by randomly masking 0–50% of the data. Inspired by masked learning [49], this design focuses on learning robust reconstruction under diverse contexts rather than emulating a specific BLE loss distribution. Figure 5 shows that real-world TWS packet loss is typically <10%; extending the range to 50% ensures robustness under more severe conditions, while also covering bursty losses, which can be viewed as localized intervals of high missingness.

Detecting the limit of cardiographic continuity refinement. Our cardiographic continuity refinement solution may not perform well when the packet loss rate is significantly high, *e.g.*, due to a long BLE link distance. We set a packet loss rate threshold τ to identify when the IMU stream is too severely corrupted. EarCardio can accordingly enter a fail-safe mode, warning the user to discard such data. We empirically determine $\tau = 24\%$ through extensive testing (Section 6.3).

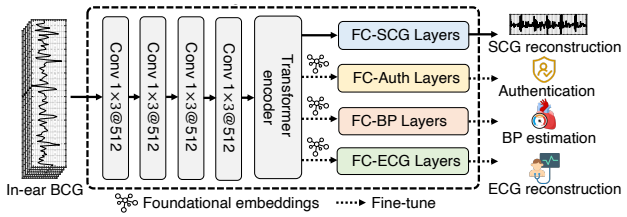


Figure 12: Cardiogram Reconstruction.

4.3 Cardiogram Reconstruction

Although our signal enhancement design improves the in-ear BCG quality, its low sampling rate and unreliable waveform limit its ability to provide detailed cardiac insights. Whereas the IMUs on COTS TWS earbuds can achieve 40 *ms* time resolution (at 25 *Hz*), reliable HRV measurement requires a time resolution of at least 10 *ms* [50]. Moreover, multi-axis in-ear BCG exhibits inconsistent peak magnitudes and waveforms due to variations in ear anatomy, vessel location, and wearing style. To overcome the challenge, we develop a super-resolution cardiogram reconstruction neural network that translates multi-axis in-ear BCG into clinically reliable SCG signals with a high sampling rate. Although sampled at 25 Hz, the target AO/AC/RF landmarks are predominantly < 10 Hz, so their timing is preserved; our model leverages cross-axis redundancy to reconstruct a higher-rate waveform.

Model design. Our model input is the output from the aforementioned denoising neural network. As shown in Figure 12, our model begins with a feature extraction stage, where 4-layer convolutional layers distills crucial spatial features from the multi-axis in-ear BCG signals. This is followed by a two-layer Transformer encoder [51], which captures long-range dependencies and sequential patterns. The encoder's attention mechanism allows the model to dynamically focus on key temporal features, particularly those associated with heartbeats. Finally, a fully connected layer decodes the processed features into single-channel SCG signals with 100 Hz sampling rate. We choose a fully connected layer instead of a Transformer decoder to enforce super-resolution upsampling while maintaining the critical timing information of the cardiographic signals by fixing the output size. The model is trained using our self-collected 25 hours of clean in-ear BCG data and their corresponding SCG ground truth (Section 5).

Cardiac-focused loss function. Recognizing that SCG reconstruction quality is primarily determined by accurately recovering heartbeat peaks and their local morphology (rather than low-energy baseline regions), we devise a *heartbeat energy-weighted loss* that explicitly biases learning toward

these informative segments. Specifically, we apply the Hilbert transform to obtain the envelope of the target SCG signal, normalize it, and use it as a weight on the MAE so that high-energy heartbeat regions contribute more to the training objective. We further combine this weighted MAE with a standard MAE, yielding a dual-objective that preserves global fidelity while sharpening peak-aligned reconstruction. The loss function is defined as:

$$\mathcal{L} = \alpha \times \left(\frac{1}{N} \sum_{i=1}^{N} |\hat{y}_i - y_i| \right) + \beta \times \left(\frac{1}{N} \sum_{i=1}^{N} w_i \, |\hat{y}_i - y_i| \right) \quad (3)$$

where N represents the number of samples, $\hat{y}_i$ is the predicted value, y_i is the ground truth value, and w_i denotes the normalized weight derived from the Hilbert-transformed heartbeat envelope. The weights α and β are hyperparameters that balance the contribution of each loss term.

Adapation to various applications using foundational cardiac-related embeddings. In-ear BCG signals can be used not only to reconstruct the SCG signals for measuring HR and HRV but also for other diverse applications (Sec. 6.5). However, each application requires collecting a large amount of paired in-ear BCG data and corresponding ground truth to train a specific model. To overcome this burden, we deliberately design the BCG-to-SCG super-resolution task as a *pretext objective*. Unlike coarse targets such as HR/HRV, SCG encodes fine-grained cardiomechanical events (*e.g.*, valve motions, systolic intervals). By requiring the encoder $f_\theta(\cdot)$ to recover SCG from noisy, low-rate BCG, we compel it to capture physiologically grounded dynamics in its latent representation $\mathbf{z} = f_\theta(x)$. Moreover, the task is formulated as a super-resolution reconstruction, which forces the encoder to learn temporal continuity and rhythmic dependencies that generalize beyond a single application. Finally, our architecture explicitly separates the encoder from the lightweight task head h_ψ, ensuring that the learned embeddings are not overfit to one task but are reusable across tasks. We therefore term $\mathbf{z}$ *foundational cardiographic embeddings*. Instead of retraining a full model, we only replace the task head h_ψ and fine-tune with few-shot paired data:

$$\hat{y} = h_\psi(\mathbf{z}), \quad \mathcal{L}_{\text{task}} = \ell(\hat{y}, y). \quad (4)$$

This design reduces the labeled data requirement by over 95% while preserving downstream usability. We empirically validate this benefit in the case studies (Sec. 6.5). *The foundation-style cardiographic pretraining provides reusable embeddings applicable diverse downstream applications.*

5 IMPLEMENTATION

Data collection. As depicted in Figure 13, our experimental setup consists of a TWS earbud paired with a smartphone. We implement apps to acquire 6-axis raw IMU data using native iOS API [52] and eSense Android library [53] for AirPods and eSense, respectively. Since these two achieve comparable performance (Sec. 3.1), our experiments use AirPods by default,

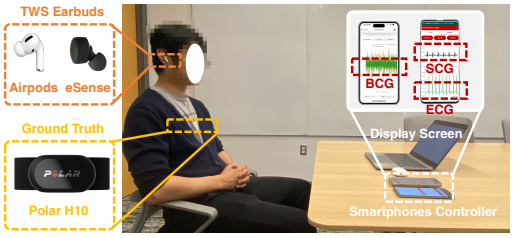


Figure 13: Experimental setup.

unless noted otherwise. All procedures were IRB-approved. We collected two datasets involving 110 participants.

1) Main dataset. We first collect clean in-ear BCG data in a controlled setting from 100 participants (40 female, 60 male), aged 19–62 with BMI ranging from 15.92 to 33.74, kg/m^2. Participants sat in a chair and wore the right earbud in the default manner, with audio functions disabled to avoid interference. Ground-truth SCG signals were recorded using the Polar H10 at a 100 Hz sampling rate, leveraging its built-in accelerometer and taking the chest-normal z-axis as the reference SCG channel. Similar to [22], the signal was then high-pass filtered at 0.5 Hz to remove respiratory components and normalized for subsequent analysis. Participants were instructed to wear the Polar H10 [54], and it was paired with a smartphone [55] for data acquisition. To ensure the generalizability of the trained models and evaluation results, we introduced variability in participants' heart rates, particularly for conditions sensitive to rate changes such as tachycardia or bradycardia. Following prior studies [56], heart rate fluctuations were stimulated by showing participant-selected videos (e.g., movies and news) on a monitor. Each participant completed a 15-minute session, yielding 25 hours of clean in-ear BCG data with about 145,000 heartbeats. This dataset was used for training, cross-validation, and evaluation of EarCardio, as well as for validating the data synthesis method.

2) Additional testing dataset. To further assess robustness under diverse conditions, we collected an additional test-only dataset from 10 participants, spanning 15 hours and approximately 87,000 heartbeats, which is used exclusively for ablation and case studies. Data were collected across different earbud sides and operating modes, as well as during motion and daily activities. Because commercial earbuds (AirPods and eSense) support IMU streaming from only one earbud at a time, we used two pairs of earbuds and smartphones, alternately enabling one earbud while disabling its paired counterpart to record data from both sides. Ground-truth SCG signals were not used in this additional dataset due to significant motion artifacts. Instead, the Polar H10 was configured to record 130 Hz ECG signals, from which precise HR and IBI measurements were derived as ground truth. In total, *we evaluated EarCardio on* 110 *participants:* 100 *in the main dataset for cross-validation, and* 10 *in the additional testing dataset for ablation and case studies.*

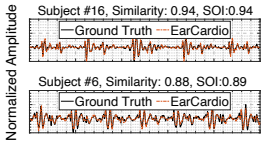


Figure 14: Samples of reconstructed SCG.

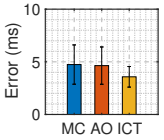


Figure 15: Features detection.

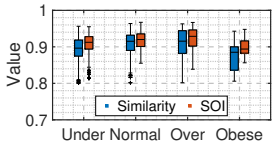


Figure 16: Impact of BMI.

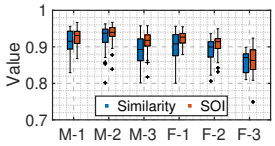


Figure 17: Age and gender.

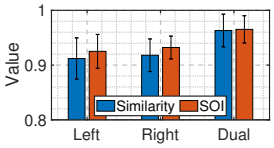


Figure 18: Side of earbuds.

Model training. The 6-axis raw IMU data and corresponding ground truth SCG signals were segmented into 5-second intervals and preprocessed with a $0.5Hz$ high-pass filter to mitigate respiratory interference. We trained our model on a server equipped with an A40 GPU over 100 epochs. The initial learning rate is set at 0.001 and reduced by half every 10 epochs. Following the data synthesis methods in Sections 4.1.2 and 4.2, we used the 25-hour clean in-ear BCG dataset from the 100 participants to synthesize a 150-hour dataset under 202 types of motion artifacts and with various packet loss rates from 0 ~ 50% for training and validation. The *main* dataset was divided into 10 groups for cross-validation, with each fold involving training on 90 participants and validation on the remaining 10.

Full implementation on smartphone. We deployed the full EarCardio pipeline, including data capture, signal processing, and neural inference, on an iPhone 15 Pro (iOS 18). Neural inference was implemented via Apple's Core ML [57].

6 EVALUATION

6.1 Evaluation Metrics

We use three metrics to assess the quality of the enhanced in-ear BCG and reconstructed cardiographic signals.

• *Cosine Similarity (Similarity)* evaluates how closely the reconstructed signal aligns with the ground truth in the time domain, with values closer to 1 indicating higher similarity.

• *Spectral Overlap Index (SOI)* measures frequency-domain accuracy, with values near 1 indicating strong spectral fidelity. It is calculated by comparing the PSD between the recovered waveform $\mathrm{PSD_e}(f)$ and the ground truth $\mathrm{PSD_g}(f)$ within the frequency range $[1Hz, 50Hz]$, defined as:

$$\mathrm{SOI} = \sum_{f=1Hz}^{50Hz} \min\left(\mathrm{PSD_e}(f), \mathrm{PSD_t}(f)\right) \Big/ \sum_{f=1Hz}^{50Hz} \mathrm{PSD_t}(f) \quad (5)$$

• *Mean Percentage Error (MPE)* quantifies the percentage error in HR and IBI estimates, two widely used parameters for CVD monitoring [58], by comparing the estimated values $V_{e,i}$ with the true values $V_{t,i}$ across N measurements.

$$\mathrm{MPE} = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{V_{e,i} - V_{t,i}}{V_{t,i}} \right| \times 100\% \quad (6)$$

Similarity and SOI are used to evaluate SCG reconstruction, which require the ground truth SCG as reference. Similarity and SOI values above 0.9 and 0.85 are considered indicative of a *very high* and *high* positive correlation for SCG reconstruction, respectively [22, 59]. MPE is used to assess the accuracy of HR and IBI measurements, with ground truth data derived from ECG signals. MPE values below 5% are used as a threshold for good accuracy in HR and IBI measurements in wearable or mobile health device studies [58].

6.2 End-to-end Performance

We first evaluate end-to-end SCG reconstruction to demonstrate generalization across 100 subjects under diverse scenarios using the *main dataset*. As shown in Figure 14, EarCardio demonstrates a *very high* positive correlation for SCG reconstruction, achieving an average Similarity of 0.92 and an SOI of 0.93. These results demonstrate *strong generalization of EarCardio across different subjects*.

Detecting key cardiac features. In SCG analysis, Mitral Valve Closure (MC) marks the onset of ventricular systole, Aortic Valve Opening (AO) indicates the start of ejection, and the Isovolumic Contraction Time (ICT) reflects left ventricular contractility [60]. Accurate detection of these events is critical for assessing cardiac mechanical function and deriving clinically relevant parameters such as stroke volume and systolic time intervals. Figure 15 shows the results of these features. EarCardio achieves mean detection errors of 4.74 ms, 4.64 ms, and 3.58 ms for MC, AO, and ICT, respectively. The average error is well within the < 10 ms range, which is considered adequate for most hemodynamic parameter estimation [61]. The results indicate that *the reconstructed SCG preserves critical cardiac diagnostic features*.

Impact of BMI. We divide the 100 subjects into four groups based on BMI: "Under" ($< 18.5kg/m^2$), "Normal" ($18.5 - 24.9kg/m^2$), "Over" ($25.0 - 29.9kg/m^2$), and "Obese" ($> 30.0kg/m^2$). As shown in Figure 16, only the obese group deviates slightly from other groups, but even this group achieves a high accuracy (Similarity: 0.87, SOI: 0.90). This suggests that *BMI has minimal impact on EarCardio*.

Impact of age and gender. We categorize the 100 subjects into six groups based on gender ("F" for female, "M" for male) and age ("1" for < 26 years, "2" for 26–45 years, "3" for > 45 years). As shown in Figure 17, *EarCardio performs consistently across both age and gender*, with average Similarity and SOI values exceeding 0.9 in most groups. However, in the F-3 group, both Similarity and SOI fall slightly below 0.9. We suspect that *age-related changes, such as reduced cardiac*

force, arterial stiffening, and muscle atrophy, likely contribute to the observed reduction in Similarity and SOI [41]. We plan to include more representative samples from this group in future studies to better understand this observation.

Impact of earbuds side. Our default model is trained on right-ear data. To study ear-side effects, we use the *additional testing dataset* with a dual-ear setup for fine-tuning, where we update only the CNN and linear projection layers for left-ear and dual-ear models while freezing the Transformer to preserve sequence modeling. Figure 18 compares the three setups: the right-ear model slightly outperforms the left, likely due to training bias, and we expect the left-ear model to improve with balanced bilateral training data. With the dual-ear setup, Similarity and SOI both increase to 0.96, showing that bilateral in-ear BCG provides complementary information and supporting the *use of dual-ear IMU streams for fine-grained cardiac sensing*. In addition, this strategy enables efficient adaptation to new earbud models with different IMU characteristics by fine-tuning only the lightweight front-end while keeping the Transformer fixed. More broadly, the pipeline can extend to other head-worn devices (e.g., clip-on earbuds or smart glasses) with minor calibration, as long as stable inertial contact is available.

6.3 Ablation Study

Resilience against low and unreliable sampling. We conducted benchmark experiments to evaluate the effectiveness of our cardiographic continuity refinement (CR) at a packet loss rate threshold of $\tau = 24\%$. This threshold was determined empirically: packet loss reaches $\tau = 24\%$ only when the earbuds are more than 10 m away from the smartphone, which corresponds to the maximum practical operating distance in typical usage scenarios. As depicted in Figure 19, CR increases the Similarity from 0.746 to 0.95, equivalent to missing only one sample within a 10-second window at 25 Hz sampling rate.

High packet loss rates and low sampling rates can lead to missing beats, as discussed in Section 3.1. To assess whether EarCardio can recover missing beats, we use the MPEs of HR and IBI as metrics. As shown in Figure 20, without EarCardio, raw in-ear BCG signals result in HR and IBI estimates with MPEs > 10% across almost all earbud operating modes. By effectively recovering missing heartbeats, CR reduces the MPEs for HR and IBI to approximately 8.1% and 6.8%, respectively, especially in Trans, PB, and ANC modes when packet loss is severe. However, denoised in-ear BCG signals still suffer from low sampling rates, which obscure key waveform features such as heartbeat peaks. Our cardiogram reconstruction (CAR) solution achieves super-resolution SCG reconstruction, further reducing HR and IBI MPEs to 2.78% and 2.71%, respectively. When combining CR and CAR, the MPEs are only 1.43% and 1.31%, demonstrating *EarCardio's effectiveness in addressing low and unreliable sampling issues*.

Resilience to motion artifacts. Next, we evaluated whether our denoising model can effectively address the motion artifacts. The motion artifact testing sets include synthetic data with 202 types of artifacts (Section 5) and real-world data collected from 10 participants across four motion states: sitting still, nodding, walking, and jogging. Each state was repeated 20 times, yielding Bp SNR (dB) distributions of (≥ 0), $(0 \sim -10)$, $(-10 \sim -20)$, and $(-20 \sim -30)$, respectively. We use the MPEs of HR and IBI as metrics as we can not capture the ground truth SCG signals during motion artifacts.

In Figure 21(a), blue dots show the mean MPE for HR and IBI under synthetic motion noise, while red dots show measurements from real-world in-ear BCG with motion noise; the two distributions closely match, indicating that the IMU simulator reproduces realistic motion-noise characteristics. Figure 21(b) further breaks down results by motion type: during sitting or head movements, MPE stays below 2%, meeting ANSI/AAMI standards [62], and remains under 10% for walking and jogging, satisfying consumer-grade requirements [58]. We further conduct an ablation study on the denoising model using both synthetic and real-world datasets (Figure 22). Without denoising, the raw in-ear BCG is overwhelmed by noise and unusable for heartbeat peak detection. SWT-based denoising reduces noise but remains insufficient. A neural denoiser (NN) alone shows promise, yet without SWT the real-data MPE only drops to 22.05% (HR) and 27.12% (IBI). Combining SWT and NN reduces errors to 4.83% for HR and 6.45% for IBI, with consistent performance across synthetic and real-world datasets, indicating that *our design can effectively handle motion artifacts*.

Robustness of cardiogram reconstruction. We compare our transformer-based cardiogram reconstruction solution with other baselines using *main dataset*, including UNet [22], LSTM [63], and the version without cardiac-focused loss (W/o CF). As shown in Figure 23, the Transformer model W/o CF already outperforms both UNet and LSTM in terms of Similarity and SOI, due to its strong ability to capture complex spatial correlations across multiple in-ear BCG channels and temporal dependencies across heartbeats. Incorporating CF further improves accuracy (Similarity: 0.92, SOI: 0.94) by directing attention to critical SCG peaks, such as the aortic valve opening and systolic ejection, which are essential for accurately deriving HRV. *While CF yields similar HR MPE, it significantly reduces IBI MPE from* 2.95% *to* 1.44%, corresponding to tens of milliseconds of error. This improvement is crucial for HRV, where small IBI errors can accumulate over time into >20% distortion in long-term metrics [64].

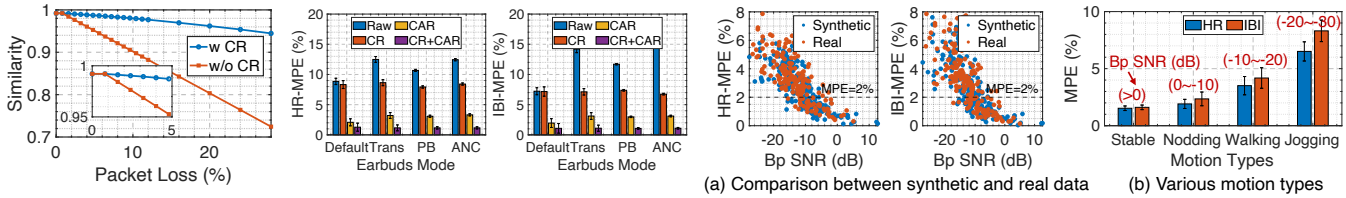


Figure 19: Efficiency of continuity refinement.

Figure 20: Performance across operation modes.

Figure 21: Performance of denoising for motion artifacts.

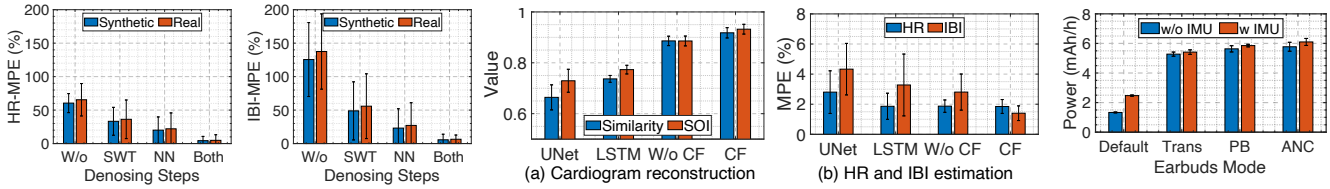


Figure 22: Denoising steps.

Figure 23: Comparison with baseline.

Figure 24: Power.

Computing Units	SWT (ms)	Denoising (ms)	Reconstruction (ms)	Total (ms)
CPU	1.45	7.46 ± 1.37	18.79 ± 1.36	27.70 ± 1.95
GPU	±0.28	9.08 ± 1.10	24.19 ± 1.50	34.72 ± 1.88
Neural	(CPU)	7.57 ± 0.28	47.00 ± 1.90	56.02 ± 1.94

Table 3: Processing time for 5-s segment on iPhone 15 Pro.

6.4 System Efficiency

Earbud Power Consumption. We measure the incremental battery overhead introduced by EarCardio, where the earbud continuously streams IMU data over BLE while the paired phone runs the full processing pipeline. Specifically, we measured the time it takes for the earbud battery to drop from 100% to 95% in Default, Trans, PB, and ANC modes using Apple AirPods Pro, with and without the BLE IMU streaming channel enabled, across 5 trials under the same mode settings. Figure 24 shows that enabling BLE IMU streaming in Default mode only increased power consumption by 1.14 mAh/h. The overhead of EarCardio is 2.6%, 3.9%, 5.7% in Trans, PB, and ANC modes, respectively, with *negligible impact on the earbud's overall power consumption*.

Smartphone Power Consumption. We measured the time it takes for the smartphone battery to drop from 100% to 95% with and without EarCardio running in the background on an iPhone 15 Pro. When using the CPU as the computational unit, the power consumption difference with and without EarCardio running in the background is only 7.26 mAh/h, which is half that of a GPS tracker (19.2mAh/h) [65].

Smartphone Processing Time. The processing time was evaluated on an iPhone 15 Pro across 100 trials using different computation units. To process segment with 5 second duration, EarCardio only requires an average 27.70 ms, 34.72 ms and 56.02 ms by using CPU, CPU + GPU and CPU + Neural Engine settings, respectively. Table 3 shows the processing time for individual modules, which demonstrates EarCardio's potential for *real-time cradiac monitoring*.

6.5 Case Study

We conducted 5 case studies to show the usage of EarCardio.

Continuous and long-term cardiac monitoring. Two participants wore earbuds for up to 150 minutes across five sessions with varying activities, postures, and earbud modes to emulate everyday use. As shown in Figure 25, EarCardio consistently achieves accurate HR estimation. Long-duration IMU-based monitoring remains drift-resistant because it avoids temporal integration, applies per-window normalization, and suppresses DC components (<0.5Hz). For longer-term evaluation, we tracked one participant over two weeks, measuring HR and HRV six times daily at 2-hour intervals using 5-minute sessions. HRV is computed via SDNN[66]. Figure 26 shows that EarCardio maintains stable HR/HRV estimates, achieving average errors of 1.68BPM (HR) and 3.68ms (HRV), closely tracking ground-truth fluctuations.

Pathological Case Detection. To illustrate EarCardio 's robustness under abnormal heart-rate regimes, we recruited three participants with tachycardia (HR > 100 BPM), bradycardia (HR < 60 BPM), and an irregular-rate case exhibiting pronounced long-term HR/HRV fluctuations. Figure 27 presents reconstructed SCG waveforms for the first two cases and shows robust IBI estimation, with MPE below 2% in 90% of segments. This suggests that *EarCardio can feasibly preserve abnormal-rate patterns in reconstructed SCG waveforms.* EarCardio 's ability to reconstruct SCG waveforms under abnormal-rate regimes stems from learning BCG-to-SCG dynamics anchored in IBI structure, rather than memorizing fixed rhythms. Our training set spans a wide HR range (48 ~ 132 BPM) induced by video stimuli (Section 5), covering both fast/slow extremes and natural IBI variability [67]. While these results indicate that EarCardio can potentially reconstruct pathological rhythms from healthy-trained data, we acknowledge the importance of further diversifying the

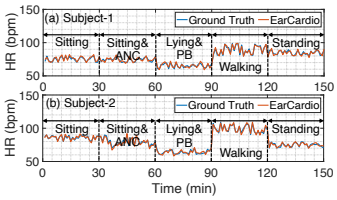


Figure 25: Continuous HR monitoring.

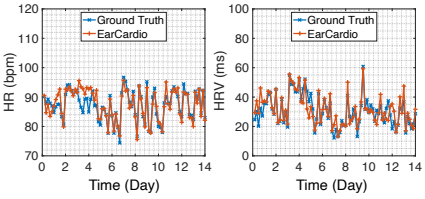


Figure 26: Long-term HR and HRV monitoring.

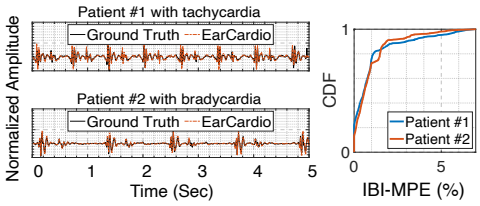


Figure 27: Analysis of pathological cases.

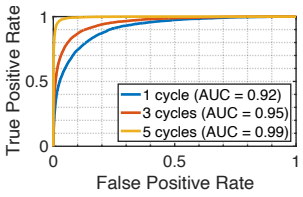


Figure 28: Authentication.

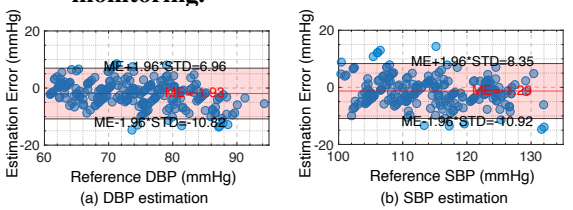


Figure 29: Blood pressure estimation.

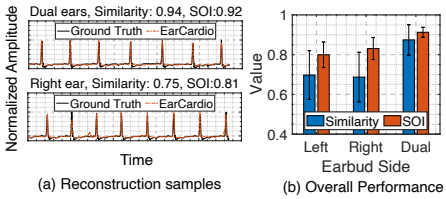


Figure 30: ECG reconstruction.

training dataset. As future work, we plan to further evaluate EarCardio on broader patient cohorts with diagnosed CVD.

User authentication. We use EarCardio for biometric authentication to enhance device security and enable personalized TWS experiences. A 5% of in-ear BCG data from 50 participants in our default dataset was used to fine-tuned the pre-trained cardiogram reconstruction model (Section 4.3) into an authentication model by adjusting the linear projection layers. We use the Area Under the Curve (AUC), a common metric for user authentication [68]. An AUC > 0.9 is considered excellent for biometric authentication [68]. As shown in Figure 28, EarCardio achieves AUCs of 0.92, 0.95 and 0.99 with 1, 3 and 5 heartbeat cycles as model input, respectively. This highlights the possibility of *using EarCardio for quick and reliable earbud-based user authentication*.

Blood pressure estimation. Previous study shows that BP can be derived by analyzing the time differences in cardiac signal propagation between multiple measurement points, *i.e.*, Pulse Transit Time (PTT) [25, 69]. Inspired by this, we leverage the dual-ear channels as natural reference points to capture PTT for BP estimation. We recruited 5 participants to collect an additional 6-hour dual-ear in-ear BCG data, while measuring the ground truth BP using an FDA-certified Checkme BP2A Monitor over one week. During data collection, each earbud streams IMU to a dedicated smartphone, and the two smartphones are precisely time-synchronized so that the dual-ear signals can be aligned for PTT extraction; each IMU/SCG window is then paired with the nearest cuff BP reading in time as the ground truth. We also fine-tune the pre-trained cardiogram reconstruction network for BP estimation in a subject-dependent setting, using 4 hours of data for training and 2 hours for testing. Figure 29 shows that EarCardio achieves strong alignment with the cuff measurements, with correlations of 0.81 for diastolic blood pressure (DBP) and 0.83 for systolic blood pressure (SBP). As a feasibility study, we will further validate the BP estimation model on a larger number of participants in future work.

Electrocardiogram reconstruction. We explore the feasibility of mapping in-ear BCG to an ECG-like waveform with EarCardio, following prior observations on mechanical-to-electrical signal relationships [9]. We collected an additional 2.5-hour dataset from 10 participants over one week using a dual-ear setup, with Polar H10 ECG recorded via chest-strap electrodes on the lower chest (sub-pectoral region) as the reference. The model was fine-tuned from the pre-trained cardiogram reconstruction network using a 2-hour training dataset and evaluated on an orthogonal 30-min testing set. We resampled the ECG signals from 130 Hz to 100 Hz to align with our model output requirements. Figure 30 illustrates an example reconstructed waveform. A single-ear setup achieves a Similarity of 0.70 and an SOI of 0.80, while the dual-ear setup improves to Similarity 0.87 and SOI 0.91, approaching high similarity (*i.e.*, ~0.9).

7 CONCLUSION

We demonstrate that IMU in commodity TWS earbuds can be repurposed for cardiac monitoring without hardware changes. EarCardio overcomes low-rate sampling and motion artifacts by reconstructing SCG signals via lightweight processing and machine learning. Extensive evaluation validates its feasibility for real-time, non-invasive cardiac monitoring in everyday earbuds, enabling future commercialization.

ACKNOWLEDGMENTS

This research was supported in part by the National Natural Science Foundation of China under Grant No. 62432004, and by a grant from the Guoqiang Institute, Tsinghua University. Xinyu Zhang was supported by an Ericsson Endowed Professorship throughout this project, and Ke Sun was supported by a Google Ph.D. Fellowship.

REFERENCES

[1] Mariachiara Di Cesare, Pablo Perel, Sean Taylor, Chodziwadziwa Kabudula, Honor Bixby, Thomas A Gaziano, Diana Vaca McGhie, Jeremiah Mwangi, Borjana Pervan, Jagat Narula, et al. The heart of the world. *Global heart*, 2024.

[2] ResearchAndMarkets.com. Global true wireless stereo earbuds market, 2022-2030: Health meets entertainment - the fusion of hearing aid functionalities in tws earbuds. Global TWS Market Report, 2023.

[3] Andrea Ferlini, Alessandro Montanari, Chulhong Min, Hongwei Li, Ugo Sassi, and Fahim Kawsar. In-ear ppg for vital signs. *IEEE Pervasive Computing*, 2021.

[4] Xiaoran Fan, Longfei Shangguan, Siddharth Rupavatharam, Yanyong Zhang, Jie Xiong, Yunfei Ma, and Richard Howard. Headfi: bringing intelligence to all headphones. In *Proceedings of ACM MobiCom*, 2021.

[5] Xiaoran Fan, David Pearl, Richard Howard, Longfei Shangguan, and Trausti Thormundsson. Apg: Audioplethysmography for cardiac monitoring in hearables. In *Proceedings of ACM MobiCom*, 2023.

[6] Tao Chen, Yongjie Yang, Xiaoran Fan, Xiuzhen Guo, Jie Xiong, and Longfei Shangguan. Exploring the feasibility of remote cardiac auscultation using earphones. In *Proceedings of ACM MobiCom*, 2024.

[7] Apple. Apple airpods, 2024. https://www.apple.com/airpods/.

[8] Fahim Kawsar, Chulhong Min, Akhil Mathur, Alessandro Montanari, Utku Günay Acer, and Marc Van den Broeck. esense: Open earable platform for human sensing. In *Proceedings of the 16th ACM Conference on Embedded Networked Sensor Systems*, pages 371–372, 2018.

[9] Lei Wang, Xingwei Wang, Dalin Zhang, Xiaolei Ma, Yong Zhang, Haipeng Dai, Chenren Xu, Zhijun Li, and Tao Gu. Knowing your heart condition anytime: User-independent ecg measurement using commercial mobile phones. *Proceedings of ACM IMWUT (UbiComp)*, 2023.

[10] Denis Jabaudon, Juan Sztajzel, Katia Sievert, Theodor Landis, and Roman Sztajzel. Usefulness of ambulatory 7-day ecg monitoring for the detection of atrial fibrillation and flutter after acute stroke and transient ischemic attack. *Stroke*, 35(7), 2004.

[11] Steven R Steinhubl, Jill Waalen, Alison M Edwards, Lauren M Ariniello, Rajesh R Mehta, Gail S Ebner, Chureen Carter, Katie Baca-Motes, Elise Felicione, Troy Sarich, et al. Effect of a home-based wearable continuous ecg monitoring patch on detection of undiagnosed atrial fibrillation: the mstops randomized clinical trial. *Jama*, 320(2), 2018.

[12] Aysha Jenea Mann. *The Analysis and Comparison of Cardiac Time Intervals via Seismocardiography*. Mississippi State University, 2024.

[13] Tommaso Sanna, Hans-Christoph Diener, Rod S Passman, Vincenzo Di Lazzaro, Richard A Bernstein, Carlos A Morillo, Marilyn Mollman Rymer, Vincent Thijs, Tyson Rogers, Frank Beckers, et al. Cryptogenic stroke and underlying atrial fibrillation. *New England Journal of Medicine*, 370(26), 2014.

[14] Apple. Apple watch, 2024. https://www.apple.com/watch/.

[15] Yandao Huang, Minghui Qiu, Lin Chen, Zhencan Peng, Qian Zhang, and Kaishun Wu. Nf-heart: A near-field non-contact continuous user authentication system via ballistocardiogram. *Proceedings of ACM IMWUT (UbiComp)*, 2023.

[16] Qiang Zhu, Xin Tian, Chau-Wai Wong, and Min Wu. Ecg reconstruction via ppg: A pilot study. In *2019 IEEE EMBS international conference on biomedical & health informatics (BHI)*. IEEE, 2019.

[17] Pritam Sarkar and Ali Etemad. Cardiogan: Attentive generative adversarial network with dual discriminators for synthesis of ecg from ppg. In *Proceedings of AAAI*, 2021.

[18] Debaditya Shome, Pritam Sarkar, and Ali Etemad. Region-disentangled diffusion model for high-fidelity ppg-to-ecg translation. In *Proceedings of AAAI*, 2024.

[19] Qiang Zhu, Xin Tian, Chau-Wai Wong, and Min Wu. Learning your heart actions from pulse: Ecg waveform reconstruction from ppg. *IEEE Internet of Things Journal*, 2021.

[20] Xin Tian, Qiang Zhu, Yuenan Li, and Min Wu. Cross-domain joint dictionary learning for ecg inference from ppg. *IEEE Internet of Things Journal*, 2022.

[21] Ella Lan. Performer: A novel ppg-to-ecg reconstruction transformer for a digital biomarker of cardiovascular disease detection. In *Proceedings of IEEE/CVF WACV*, 2023.

[22] Unsoo Ha, Salah Assana, and Fadel Adib. Contactless seismocardiography via deep learning radars. In *Proceedings of the 26th annual international conference on mobile computing and networking*, pages 1–14, 2020.

[23] Alistair EW Johnson, Tom J Pollard, Lu Shen, Li-wei H Lehman, Mengling Feng, Mohammad Ghassemi, Benjamin Moody, Peter Szolovits, Leo Anthony Celi, and Roger G Mark. Mimic-iii, a freely accessible critical care database. *Scientific data*, 3(1):1–9, 2016.

[24] Ary L Goldberger, Luis AN Amaral, Leon Glass, Jeffrey M Hausdorff, Plamen Ch Ivanov, Roger G Mark, Joseph E Mietus, George B Moody, Chung-Kang Peng, and H Eugene Stanley. Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. *circulation*, 2000.

[25] Ananta Narayanan Balaji, Andrea Ferlini, Fahim Kawsar, and Alessandro Montanari. Stereo-bp: Non-invasive blood pressure sensing with earables. In *Proceedings of ACM HotMobile*, 2023.

[26] Pan Qian, Eddie Siahaan, Scott C. Grinker, and Jean-Jacques LeBlanc. Earbuds with biometric sensing, March 2017.

[27] Yetong Cao, Chao Cai, Fan Li, Zhe Chen, and Jun Luo. Enabling passive user authentication via heart sounds on in-ear microphones. *IEEE Transactions on Dependable and Secure Computing*, 2024.

[28] Yetong Cao, Chao Cai, Fan Li, Zhe Chen, and Jun Luo. Heartprint: Passive heart sounds authentication exploiting in-ear microphones. In *IEEE INFOCOM 2023-IEEE Conference on Computer Communications*, pages 1–10. IEEE, 2023.

[29] Kayla-Jade Butkow, Ting Dang, Andrea Ferlini, Dong Ma, Yang Liu, and Cecilia Mascolo. An evaluation of heart rate monitoring with in-ear microphones under motion. *Pervasive and Mobile Computing*, 100:101913, 2024.

[30] Kayla-Jade Butkow, Ting Dang, Andrea Ferlini, Dong Ma, and Cecilia Mascolo. heart: Motion-resilient heart rate monitoring with in-ear microphones. In *2023 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pages 200–209, 2023.

[31] Yetong Cao, Chao Cai, Anbo Yu, Fan Li, and Jun Luo. Earace: Empowering versatile acoustic sensing via earable active noise cancellation platform. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 7(2):1–23, 2023.

[32] Tianfang Zhang, Cong Shi, Payton Walker, Zhengkun Ye, Yan Wang, Nitesh Saxena, and Yingying Chen. Passive vital sign monitoring via facial vibrations leveraging ar/vr headsets. In *Proceedings of ACM MobiSys*, 2023.

[33] Eric S Winokur, David Da He, and Charles G Sodini. A wearable vital signs monitor at the ear for continuous heart rate and pulse transit time measurements. In *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2012.

[34] Bo Håkansson, Fausto Woelflin, Anders Tjellström, and William Hodgetts. The mechanical impedance of the human skull via direct bone conduction implants. *Medical Devices: Evidence and Research*, pages 293–313, 2020.

[35] Hao Zhou, Md Mahbubur Rahman, Mehrab Bin Morshed, Yunzhi Li, Md Saiful Islam, Larry Zhang, Jungmok Bae, Christina Rosa, Wendy Berry Mendes, and Jilong Kuang. Know your heart better: Multimodal cardiac output monitoring using earbuds. In *ICASSP 2025*

- *2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1–5, 2025.

[36] Md Saiful Islam, Md Mahbubur Rahman, Mehrab Bin Morshed, David J Lin, Yunzhi Li, Hao Zhou, Wendy Berry Mendes, and Jilong Kuang. Ballistobud: Heart rate variability monitoring using earbud accelerometry for stress assessment. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, CHI '25, New York, NY, USA, 2025. Association for Computing Machinery.

[37] Chang-Sei Kim, Stephanie L Ober, M Sean McMurtry, Barry A Finegan, Omer T Inan, Ramakrishna Mukkamala, and Jin-Oh Hahn. Ballistocardiogram: Mechanism and potential for unobtrusive cardiovascular health monitoring. *Scientific reports*, 2016.

[38] George A McKay. Signal-to-noise ratio in the bandpass output of a discriminator. *IEEE Transactions on Aerospace and Electronic Systems*, 1970.

[39] Apple. Choose your airpods pro ear tips and use the ear tip fit test, 2024. https://support.apple.com/en-us/119849.

[40] Cambridge Nokia Bell Labs. esense user documentation, 2019. https://www.esense.io/share/eSense-User-Documentation.pdf.

[41] Ibrahim Sadek, Jit Biswas, and Bessam Abdulrazak. Ballistocardiogram signal processing: A review. *Health information science and systems*, 2019.

[42] Vishal Varun Tipparaju, Kyle R Mallires, Di Wang, Francis Tsow, and Xiaojun Xian. Mitigation of data packet loss in bluetooth low energy-based wearable healthcare ecosystem. *Biosensors*, 2021.

[43] Shuto Nagai, Daisuke Anzai, and Jianqing Wang. Motion artefact removals for wearable ecg using stationary wavelet transform. *Healthcare technology letters*, 2017.

[44] Emilio J Ochoa and Luis C Revilla. Signal filtering and peak analysis of ballistocardiography for heartbeat detection. In *International Conference on Biomedical and Health Informatics*. Springer, 2022.

[45] Dongyeol Seok, Sanghyun Lee, Minjae Kim, Jaeouk Cho, and Chul Kim. Motion artifact removal techniques for wearable eeg and ppg sensor systems. *Frontiers in Electronics*, 2021.

[46] Yinghao Huang, Manuel Kaufmann, Emre Aksan, Michael J Black, Otmar Hilliges, and Gerard Pons-Moll. Deep inertial poser: Learning to reconstruct human pose from sparse inertial measurements in real time. *ACM Transactions on Graphics (TOG)*, 2018.

[47] Naureen Mahmood, Nima Ghorbani, Nikolaus F Troje, Gerard Pons-Moll, and Michael J Black. Amass: Archive of motion capture as surface shapes. In *Proceedings of ICCV*, 2019.

[48] Omer T Inan, Pierre-Francois Migeotte, Kwang-Suk Park, Mozziyar Etemadi, Kouhyar Tavakolian, Ramon Casanella, John Zanetti, Jens Tank, Irina Funtova, G Kim Prisk, et al. Ballistocardiography and seismocardiography: A review of recent advances. *IEEE journal of biomedical and health informatics*, 2014.

[49] Kaiming He, Xinlei Chen, Saining Xie, Yanghao Li, Piotr Dollár, and Ross Girshick. Masked autoencoders are scalable vision learners. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 16000–16009, 2022.

[50] Brian W Johnston, Richard Barrett-Jolley, Anton Krige, and Ingeborg D Welters. Heart rate variability: Measurement and emerging use in critical care medicine. *Journal of the Intensive Care Society*, 2020.

[51] A Vaswani. Attention is all you need. *Advances in Neural Information Processing Systems*, 2017.

[52] Apple. Apple developer cmheadphonemotionmanager, 2024. https://developer.apple.com/documentation/coremotion/cmheadphonemotionmanager.

[53] eSense. esense android library, 2024. https://github.com/pervasive-systems/eSense-Android-Library.

[54] Polor. Polor h10 user manual, 2024. https://support.polar.com/e_manuals/h10-heart-rate-sensor/polar-h10-user-manual-english/manual.pdf.

[55] Polor. Polor flow app, 2024. https://support.polar.com/en/support/Flow_app.

[56] Daniel C Richardson, Nicole K Griffin, Lara Zaki, Auburn Stephenson, Jiachen Yan, Thomas Curry, Richard Noble, John Hogan, Jeremy I Skipper, and Joseph T Devlin. Engagement in video and audio narratives: Contrasting self-report and physiological measures. *Scientific Reports*, 2020.

[57] Apple. Apple developer core ml, 2024. https://developer.apple.com/documentation/coreml/.

[58] Benjamin W Nelson and Nicholas B Allen. Accuracy of consumer wearable heart rate measurement during an ecologically valid 24-hour period: intraindividual validation study. *JMIR mHealth and uHealth*, 7(3):e10828, 2019.

[59] Mavuto M Mukaka. A guide to appropriate use of correlation coefficient in medical research. *Malawi medical journal*, 24(3):69–71, 2012.

[60] Tilendra Choudhary, LN Sharma, and Manas Kamal Bhuyan. Automatic detection of aortic valve opening using seismocardiography in healthy individuals. *IEEE journal of biomedical and health informatics*, 23(3):1032–1040, 2018.

[61] Ghufran Shafiq, Sivanagaraja Tatinati, Wei Tech Ang, and Kalyana C Veluvolu. Automatic identification of systolic time intervals in seismocardiogram. *Scientific reports*, 6(1):37524, 2016.

[62] Francesco Sartor, Jos Gelissen, Ralph Van Dinther, David Roovers, Gabriele B Papini, and Giuseppe Coppola. Wrist-worn optical and chest strap heart rate comparison in a heterogeneous sample of healthy individuals and in coronary artery disease patients. *BMC Sports Science, Medicine and Rehabilitation*, 10(1):10, 2018.

[63] Rashmi Kumari, Surita Sarkar, Debeshi Dutta, Pabitra Das, and Amit Acharyya. P2e-lgan: Ppg to ecg reconstruction methodology using lstm based generative adversarial network. In *2024 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, 2024.

[64] Junichiro Hayano and Emi Yuda. Pitfalls of assessment of autonomic function by heart rate variability. *Journal of physiological anthropology*, 2019.

[65] Rahul Murmuria, Jeffrey Medsger, Angelos Stavrou, and Jeffrey M Voas. Mobile application and device power usage measurements. In *2012 IEEE sixth international conference on software security and reliability*. IEEE, 2012.

[66] Fred Shaffer and Jay P Ginsberg. An overview of heart rate variability metrics and norms. *Frontiers in public health*, 2017.

[67] Angel Moya, Richard Sutton, et al. Guidelines for the diagnosis and management of syncope (version 2009): the task force for the diagnosis and management of syncope of the european society of cardiology (esc). *European heart journal*, 30(21):2631–2671, 2009.

[68] Chae Lin Seok, Young Do Song, Byeong Seon An, and Eui Chul Lee. Photoplethysmogram biometric authentication using a 1d siamese network. *Sensors*, 2023.

[69] Yetong Cao, Shujie Zhang, Fan Li, Zhe Chen, and Jun Luo. hbp-fi: Contactless blood pressure monitoring via deep-analyzed hemodynamics. In *IEEE INFOCOM 2024-IEEE Conference on Computer Communications*, pages 1211–1220. IEEE, 2024.