Knowing Your Heart Condition Anytime: User-Independent ECG Measurement Using Commercial Mobile Phones

LEI WANG, Soochow University University, China
XINGWEI WANG, Shenzhen Institutes of Advanced Technology, CAS, China
DALIN ZHANG, Aalborg University, Denmark
XIAOLEI MA, Soochow University, China
YONG ZHANG, Shenzhen Institutes of Advanced Technology, CAS, China
HAIPENG DAI, Nanjing University , China
CHENREN XU, Peking University, China
ZHIJUN LI, Soochow University University, China
TAO GU, Macquarie University , Australia

Electrocardiogram (ECG) monitoring has been widely explored in detecting and diagnosing cardiovascular diseases due to its accuracy, simplicity, and sensitivity. However, medical- or commercial-grade ECG monitoring devices can be costly for people who want to monitor their ECG on a daily basis. These devices typically require several electrodes to be attached to the human body which is inconvenient for continuous monitoring. To enable low-cost measurement of ECG signals with off-the-shelf devices on a daily basis, in this paper, we propose a novel ECG sensing system that uses acceleration data collected from a smartphone. Our system offers several advantages over previous systems, including low cost, ease of use, location and user independence, and high accuracy. We design a two-tiered denoising process, comprising SWT and *Soft-Thresholding*, to effectively eliminate interference caused by respiration, body, and hand movements. Finally, we develop a multi-level deep learning recovery model to achieve efficient, real-time and user-independent ECG measurement on commercial mobile phones. We conduct extensive experiments with 30 participants (with nearly 36,000 heartbeat samples) under a user-independent scenario. The average errors of the PR interval, QRS interval, QT interval, and RR interval are 12.02 ms, 16.9 ms, 16.64 ms, and 1.84 ms, respectively. As a case study, we also demonstrate the strong capability of our system in signal recovery for patients with common heart diseases, including tachycardia, bradycardia, arrhythmia, unstable angina, and myocardial infarction.

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

Additional Key Words and Phrases: Heartbeat, Seismocardiography.

ACM Reference Format:
Lei Wang, Xingwei Wang, Dalin Zhang, Xiaolei Ma, Yong Zhang, Haipeng Dai, Chenren Xu, Zhijun Li, and Tao Gu. 2023. Knowing Your Heart Condition Anytime: User-Independent ECG Measurement Using Commercial Mobile Phones. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 7, 3, Article 131 (September 2023), 28 pages. https://doi.org/10.1145/3610871

Corresponding author: Dalin Zhang.
Authors' addresses: Lei Wang, Soochow University University, China, Email:wanglei@suda.edu.cn; Xingwei Wang, Shenzhen Institutes of Advanced Technology, CAS, China; Dalin Zhang, Aalborg University, Denmark; Xiaolei Ma, Soochow University, China; Yong Zhang, Shenzhen Institutes of Advanced Technology, CAS, China; Haipeng Dai, Nanjing University , China; Chenren Xu, School of Computer Science, School of Electronics Engineering and Computer Science, Peking University, Beijing, China; Zhijun Li, Soochow University University, China; Tao Gu, Macquarie University , Australia.

2474-9567/2023/9-ART131 $15.00
https://doi.org/10.1145/3610871

(a) Typical application scenario with our system

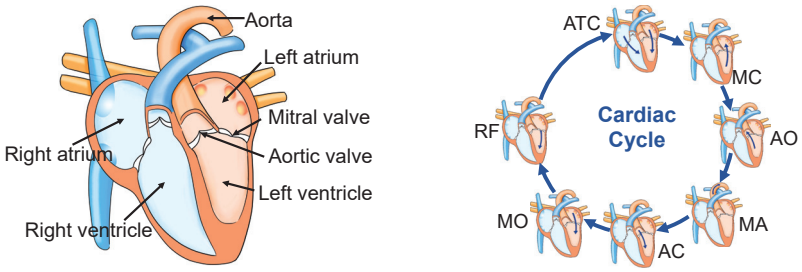
(b) Interface of our Android APP.

Fig. 1. ECG recovery system.

1 INTRODUCTION

According to the World Health Organization [1], cardiovascular diseases (CVD) are the leading cause of death globally, claiming an estimated 17.9 million lives each year [2–4]. Cardiovascular diseases include coronary heart disease, acute myocardial infarction (a.k.a. heart attack), and arrhythmia. In these diseases, sudden cardiac arrest can often occur without warning signs. Therefore, taking preventative measures is crucial for early diagnosis and treatment of CVD. Various wearable systems, such as fitness trackers [5–8], chest straps [9–11], and smartwatches[12–16], are available today that can monitor different aspects of the heart, such as heart rate and heart rate variability (HRV). However, heart rate and its variability are merely reference indicators and they cannot be used as definitive proof of the presence of heart disease. In contrast, Electrocardiogram (ECG) signals [17–19] directly measure the heart's electrical activity and can detect minor abnormalities in the heart. ECG has been widely used as an accurate, simple, and sensitive method for detecting and diagnosing cardiac diseases.

Commercial ECG devices can be broadly categorized as medical-grade and commercial-grade. Medical-grade ECG devices offer various functions such as diagnosis, report generation, and data storage. These devices are very costly, ranging from $10,000-30,000, and are generally used in clinical settings. Therefore, they may not be suitable for daily use by the general public [20]. On the other hand, commercial-grade ECG devices cost a few hundred dollars [21]. However, these devices typically require attaching several electrodes to the human body, making them less convenient for daily use. Although electrodes can be integrated into wearable devices such as smartwatches [22–24], their performance highly depends on the precise location and proper contact of the electrode. There have been attempts to convert heartbeat signals into ECG signals using commercial sensing modalities such as ultra-wideband (UWB) [25], millimeter wave (mmWave) [26], and geophones [27]. However, these fixed devices only work when users are located in a particular area, making it challenging to achieve location-independent (i.e., anywhere) daily heartbeat monitoring. A recent solution proposed by Cao et al. [28] approximates the ECG signal through the detection of wrist pulse using a smartwatch. Although the solution is location independent, it relies on specialized equipment and can suffer from poor accuracy when reconstructing ECG signals for unseen users. In this paper, we propose a novel low-cost ECG sensing system that accurately measures ECG signals using acceleration data, i.e., Seismocardiogram (SCG) signals, collected from mobile devices. Our system offers several advantages over previous systems, including low-cost, ease of use, location independence, user independence, and high accuracy. Fig. 1(a) illustrates a typical scenario where a patient can measure their ECG readings (Fig. 1(b)) by placing a mobile phone against the chest. The signals transmitted to a cloud server allow physicians to diagnose heart-related conditions remotely and in real-time.



(a) The structure of the heart (b) cardiac cycle

Fig. 2. Cardiac structure and cycle.

Although promising, there are various challenges that need to be addressed. The first challenge is noise interference, where the accelerometer captures not only the heartbeat signal, but also data related to breathing, hand movements, and system noise. Consequently, contaminated heartbeat signals have to be cleaned prior to recovery. To address this challenge, we propose a two-step strategy: 1) eradicating the effects of respiration and hand movements using the Stationary Wavelet Transform Thresholding (*SWT-Thresholding*) method, and 2) eliminating system noise by the *Soft-Thresholding* method. The second challenge is recovering both coarse-grained and fine-grained ECG characteristics. Simple models like autoencoder (AE) fall short of capturing fine-grained details. We propose a Multi-level ECG-recovery Model (MEM) that can successfully embed low-, middle-, and high-level features in the encoder and eventually recover them to ECG with full details. To extract the finest-grained high-level features from SCG, we extend the vigorous Atrous Spatial Pyramid Pooling (ASPP) module to one-dimensional (1D) data. Additionally, we use a Huber loss to train the MEM, which can balance convergence speed and robustness to noise effectively.

In summary, this paper makes the following contributions.

• To the best of our knowledge, we are the first to develop a low-cost ECG sensing system that utilizes acceleration data from commercially available mobile phones. This system offers several advantages such as user and location independence, as well as easy accessibility. We successfully demonstrate a promising mobile application that enables users to obtain real-time ECG readings by placing their mobile phones against the chest. The system allows physicians to remotely diagnose heart-related conditions, thereby enhancing the feasibility of early-stage disease screening.

• We design a two-tiered denoising process, comprising SWT and *Soft-Thresholding*, that effectively eliminates interference caused by respiration, body, and hand movements. To achieve efficient and user-independent ECG measurement, we develop a multi-level deep learning recovery model which can embed both coarse- and fine-grained details of SCG fluctuations and recover ECG signals with full details.

• We implement the real-time ECG recovery system on a commercial mobile phone and conduct extensive experiments with 30 participants (with nearly 36,000 heartbeat samples) in a user-independent scenario. The results show that the average errors of the PR interval, QRS interval, QT interval, and RR interval are 12.02 ms, 16.9 ms, 16.64 ms, and 1.84 ms, respectively. In our case study, we demonstrate the strong capability of this system in recovering signals for patients with common heart diseases, such as tachycardia, bradycardia, arrhythmia, unstable angina, and myocardial infarction.

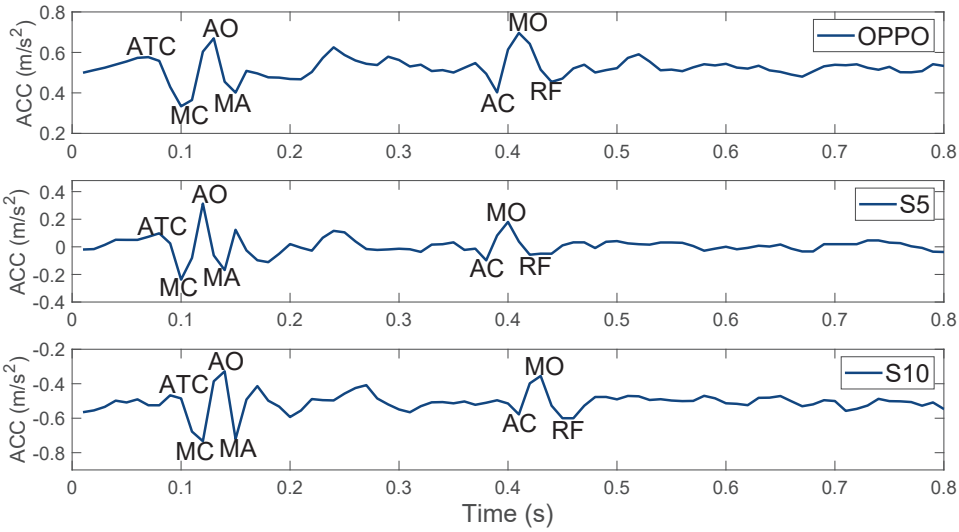


Fig. 3. Accelerometer readings from three types of mobile phones.

2 UNDERSTANDING SCG SIGNALS

2.1 SCG Signal and Cardiac Cycle

Similar to ECG signals which provide clinical information, including five peaks i.e., P, Q, R, S, T, to form the basis of measuring the cardiac activity [29, 30], SCG signals collected by the accelerometer are also effective in capturing heartbeat activity information. Therefore, SCG measurements are often used to assist ECG measurement in the diagnosis of various cardiac disorders [31, 32]. Heartbeat activity characterizes the three-dimensional heart deformation caused by the stimulation of cardiac muscle [33]. As shown in Fig. 2(a), the human heart consists of two upper atria and two lower ventricles [34]. A full heartbeat motion cycle goes through the following seven mechanical events caused by continuous cardiac systole and diastole [35–37], as shown in Fig. 2(b). (i) Atrial contraction (ATC): The electrical depolarization of the atria begins this stage of atrial muscle contraction. With the atria contraction, the pressure within the atria increases, forcing more blood to flow through the open atrioventricular valves and causing a rapid flow of blood into the ventricles. (ii) Mitral valve closure (MC): This stage implies ventricular depolarization, resulting in mitral valve closure and a rapid increase in pressure inside the ventricular. (iii) Aortic valve opening (AO): At this stage, blood rushes into the aorta from the left and right ventricles. Ejection begins when the pressure inside the ventricle exceeds the pressure inside the aorta, causing the aortic valve to open. (iv) Point of maximal blood acceleration in the aorta (MA): In this stage, the blood flows from the left ventricle to the aorta with maximal acceleration. (v) Aortic valve closure (AC): Sudden closure of the aortic valve causes isovolumic relaxation. The valve closure is associated with a small amount of blood returning to the ventricle and a characteristic notch in the aorta. (vi) Mitral valve opening (MO): During the left ventricular diastole, blood flows from the left atrium to the left ventricle as the left ventricular myocardium diastole and the mitral valve opens after a decrease in left ventricular pressure. The filling is due to the diastole of the ventricular myocardium, leading to a pressure gradient that allows blood to flow rapidly from the left atrium over the mitral valve. (vii) Rapid filling of the left ventricle (RF): This stage occurs when the pressure of the left atrial exceeds that of the left ventricular, and the mitral valve opens, allowing passive blood to flow into the left ventricle.

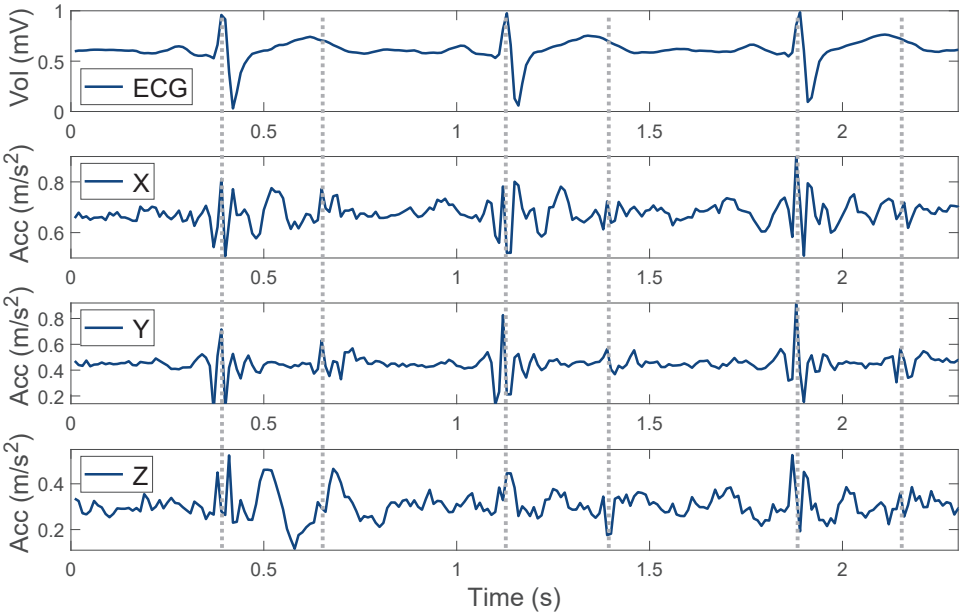


Fig. 4. Comparison between SCG and ECG signals.

These cardiac events in one heartbeat cycle can be easily captured and identified using a miniaturized accelerometer on the human sternum [36]. Inspired by this, we press three types of mobile phones vertically on the same chest, respectively, while the user holds his/her breath, as shown in Fig. 1(a). Then, we record the acceleration data of the Y-axis direction (i.e., the readings from the bottom pointing to the top of the phone) in Fig. 3. It shows that each motion stage can be identified, and the patterns of the waveform from the readings of these phones are similar, demonstrating that the accelerometer readings provided by commercial mobile phones can be used for capturing SCG.

2.2 Relationship between SCG and ECG

Although SCG signals can be used to characterize cardiac activity, current medical analysis and diagnosis of heart disease still rely on ECG. In this light, we focus on the in-depth "translation" of SCG signals collected from commercial mobile phones into ECG signals. In this section, we discuss the intrinsic relationship between SCG signals and ECG signals to illustrate the feasibility of recovering SCG to ECG waveform.

Although the SCG signals describe the mechanical vibration of the sternum and the ECG signals measure the corresponding electrical activity, they essentially characterize the same heartbeat activity resulting from periodical cardiac muscle systole and diastole. Therefore, we believe that there is a strong correlation between these two time-series waveforms. To verify this point, we invited a volunteer to press the mobile phone in front of the sternum, as shown in Fig. 1(a) and collected the accelerometer signal for 2.3 seconds, which corresponds to 230 sampling points, given that the sampling rate is fixed at 100 Hz. At the same time, we recorded the ECG measurements as a reference signal. Fig. 4 shows that the SCG signal depicts the same systolic and diastolic heartbeat activity as the ECG signal in each cardiac cycle. The P wave in the ECG signal corresponds to the ATC state in the SCG signal, implying atrial systole. Upon that, the QRS wave in the ECG signal corresponds to the range of states from MC to MA in the SCG signal, indicating ventricular systole. Finally, the T wave in the ECG signal is associated with

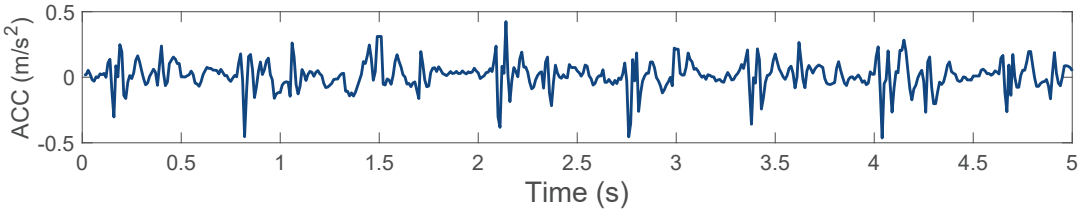
the range of states from AC to RF in the SCG signal, which indicates ventricular diastole. Based on the above observations, we can conclude that SCG signals collected from mobile phones and ECG signals are inherently correlated, demonstrating the great feasibility of "translating" SCG signals from mobile phones to ECG signals. In addition, although the accelerometer readings show periodic patterns in the X, Y, and Z axes, the signal-to-noise ratio (SNR) in the Y-axis is significantly higher than that in the other two axes because the phone is pressed vertically against the chest. Therefore, in the default scenario, our system employs only the accelerometer readings in the Y direction to reduce noise interference.

3 TWO-TIERED SIGNAL PURIFICATION

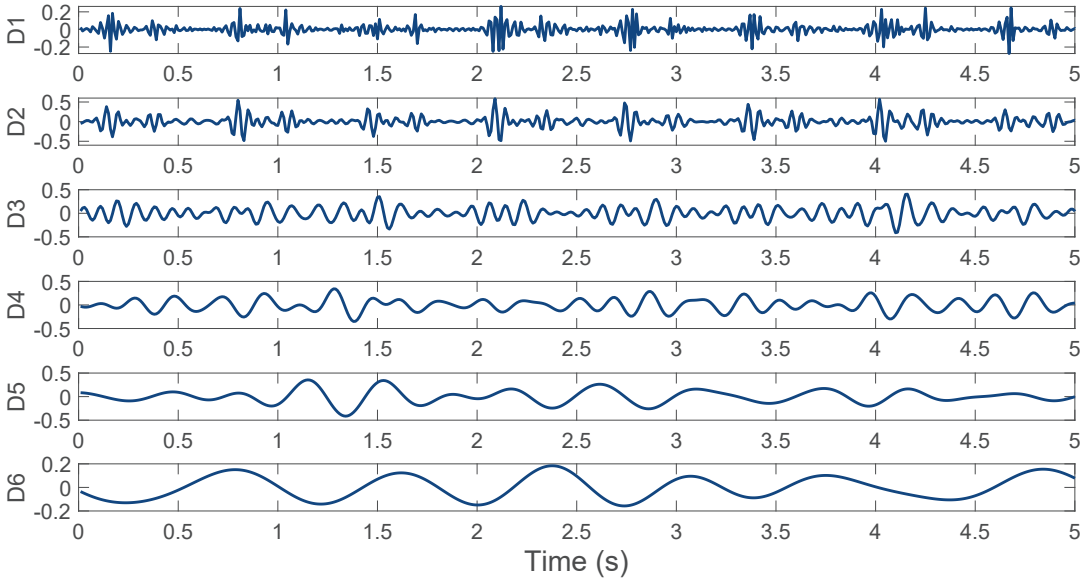
Much effort has been devoted to removing noise from biological signals. These studies can be divided into the following categories: high-pass filters [38] or low-pass filters [39, 40], notch filters [41], empirical mode decomposition (EMD) [42, 43], and adaptive filters [44]. Although high-pass or low-pass filters can remove interference noise, some key features would also be eliminated. Likewise, the notch filters also filter out signals outside the specific frequency range, thus affecting the integrity of a signal. Another commonly used technique is EMD, which decomposes the signal into a series of frequency- and amplitude-modulated components to enable denoising. However, the low-frequency components are also wiped off, leading to the loss of useful information. Consequently, the challenge lies in efficiently extracting the SCG signal induced by heartbeat movement using the accelerometer data collected from mobile phones. In this section, we introduce a two-tiered denoising scheme specifically designed to eliminate various forms of noise, such as motion-induced signals and other residual noise.

3.1 Motion-induced Noise Removal

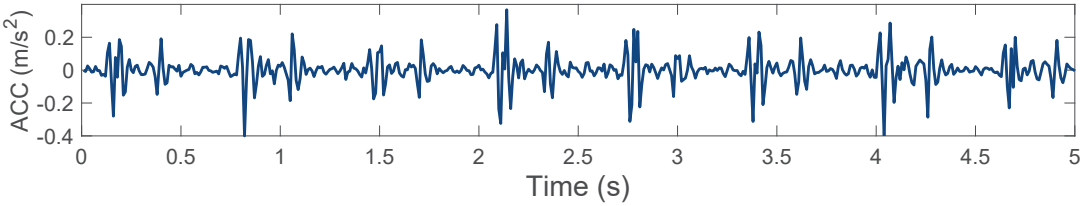
Recent studies [28, 45] have demonstrated the efficiency of Stationary Wavelet Transform (SWT) in mitigating artifact-induced noise from heart-induced signals due to its superiority of struggling against the shifts of signals in the time domain. We are inspired to use the SWT-based technique to extract the components related to cardiac movements from the accelerometer data contaminated with motion-induced signals from respiration, body, and hand movements. In detail, it first decomposes raw signals in Fig. 5(a) to derive the wavelet coefficients sequence and scaling coefficients sequence by performing the SWT operation. Specifically, the wavelet coefficient sequence and scaling coefficients sequence can be represented as $\{d_1, d_2, \ldots, d_M\}$ and $\{a_1, a_2, \ldots, a_M\}$, respectively, where M is denoted as the level of decomposition. Here, we select the Coif5 as the wavelet due to its superiority in time-frequency localization and smoothness. Given that the frequency range of heartbeat components falls within $0.8 \sim 50$ Hz, we set M to 6, based on a sampling rate of 100 Hz. As a result, levels 1 to 6 represent signal components within the frequency ranges of $25 \sim 50$ Hz, $12.5 \sim 25$ Hz, $6.25 \sim 12.5$ Hz, $3.13 \sim 6.25$ Hz, $1.56 \sim 3.13$ Hz, and $0.78 \sim 1.56$ Hz, respectively. Since respiration falls within the frequency range of $0.2 \sim 0.33$ Hz [46], we can effectively eliminate the impact of respiration during the decomposition step. Fig. 5(b) shows the corresponding SWT output of wavelet coefficient sequences. Second, we focus on removing the noise from body/hand movements. The basic idea behind this step is that the amplitude of the vibration caused by body movements is much greater than the vibration in response to the heartbeat. Specifically, in each wavelet coefficient sequence, we first calculate the energy within a period of 0.1 s and then locate all the peaks (local maximum points) $P_m(i)$, where i indicates the order in time. We then calculate the mean μ_m and standard deviation δ_m of all peaks in each level and set the empirical threshold Thr_m to $\mu_m + 2\delta_m$. If the peak $P_m(i)$ is higher than Thr_m, the peak will be identified as the body or hand-induced activity. Third, the waveform related to heart-induced activity exhibits a quasi-periodic pattern, unlike those associated with minor noisy movements. This distinction can be effectively leveraged to identify and isolate such noise. In detail, we denote the interval between two adjacent peaks, e.g., $P_m(i)$ and $P_m(i+1)$, as $\tau_m(i)$. Similarly, the empirical threshold of interval $\tau_m^{Thr}(i)$ is calculated as the median value of $\tau_m(i)$. Since the peak induced by the heartbeat should appear periodically in the SCG signals, $P_m(i)$ is determined within the heartbeat cycle when there are two consecutive time adjacent intervals (i.e., $\tau_m(i)$, $\tau_m(i+1)$) that are larger than $\tau_m^{Thr}(i)$.



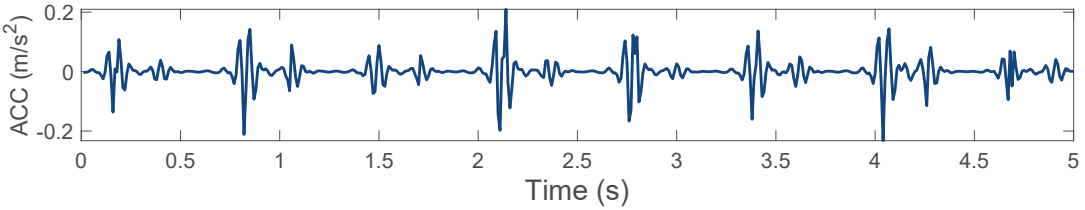
(a) Raw SCG signals



(b) Output of SWT



(c) Recovered SCG signal after SWT



(d) Clean SCG signal after Soft-Thresholding

Fig. 5. Illustration of Denoising Process.

Otherwise, $P_m(i)$ is identified as the noise induced by other tiny movements. By locating the start and end points of the heart-induced peaks, we can separate the heart-induced segment and noisy segment in each level of wavelet coefficients. Upon that, we replace the noisy time segment with zero. Thereafter, by performing the inverse SWT, we can recover the purified SCG signals. Fig. 5(c) shows that the influence of the artificial movements can be significantly reduced and the heartbeat pattern of the purified signal is much clearer than that of the original signal. Note that we can clearly observe that the SCG signal's amplitude partially decreases following the SWT-based process. This is primarily because the SWT-based method retains components within the frequency range of 0.78 to 50 Hz, while the parts outside this range are classified as noise and hence discarded. This reduction in heartbeat amplitude information will inevitably influence our system's recovery capability to a certain degree. However, since the proportion of the signal that falls outside this frequency range is relatively minor and often contaminated by high-frequency noise, for instance, random system noise, the system's overall impact from this direct removal is marginal.

3.2 Small Noise Reduction

Although the SWT-based technique is highly effective in removing most of the motion-induced noise, some residual movement noise may still remain in the wavelet coefficient sequences of each layer. This is due to the heartbeat's pseudo-periodic nature, which could potentially have a destructive effect on ECG recovery. Therefore, an additional purification step is required to enhance the quality of the SCG signals. Recent works have demonstrated the feasibility of using the adaptive filter to remove this residual noise [47, 48]. However, this technique is constrained by the following limitations. First, it necessitates a reference signal, either from another sensor or the original signal which is not feasible for practical applications. Second, the adaptive filter employs an adaptive algorithm (e.g., Least Mean Squares [47, 48]), adjusting its parameters at every time step, to efficiently respond to signal variations. However, this process necessitates a significant expenditure of computational resources.

In this paper, we use the *Soft-Thresholding* algorithm [49–51] to design a new denoising scheme based on wavelet coefficient sequences to further facilitate the extraction of clear vibrations of heartbeat motion. It can reconstruct unknown waveforms from data containing independent and identical Gaussian distributions. Our *Soft-Thresholding* based denoising scheme undergoes the following two steps. First, it calculates the empirical global threshold [50] which can be represented as: $\lambda = \hat{\sigma}\sqrt{2\ln(N)}$, where N is the length of data and $\hat{\sigma}$ indicates the noise level. Second, to derive the filtered wavelet coefficient, we apply the soft thresholding on each element of the wavelet coefficient based on the chosen global threshold as: $v_m(n) = \text{sign}(|d_m(n)| - \lambda)_+$, where x_+ is defined as x when $x \geq 0$ and otherwise 0. Then, by performing inverse SWT on the filtered result, the random noise is significantly weakened, and we can see the clearer heartbeat pattern, as shown in Fig. 5(d).

4 ECG MEASUREMENT

4.1 Data Normalization

When the mobile phone's built-in accelerometer captures the SCG signal, our system will leverage the characteristics of its waveform to recover the ECG signal through our deep learning model (details in Section 4.1). However, several factors may influence the accuracy of ECG recovery owing to the lack of data diversity, e.g., body posture, pressing strength. Therefore, we appeal to data normalization so that the data can cover most potential scenarios.

Users may want to use their mobile phones in different ways, e.g., with different postures or pressing strengths. To simulate these scenarios, we first ask a subject to keep two different postures, sitting and lying, with the phone always perpendicular to the chest. Then, the subject was asked to hold the phone toward his chest with different intensities, i.e., light and hard press. Although the heartbeat waveform pattern of the subject is consistent in both scenarios, their absolute accelerations and scaling range differ significantly due to the effect of the gravity [36] and the amount of pressure, respectively. To reduce the variations of the SCG signals, we employ the Min-Max

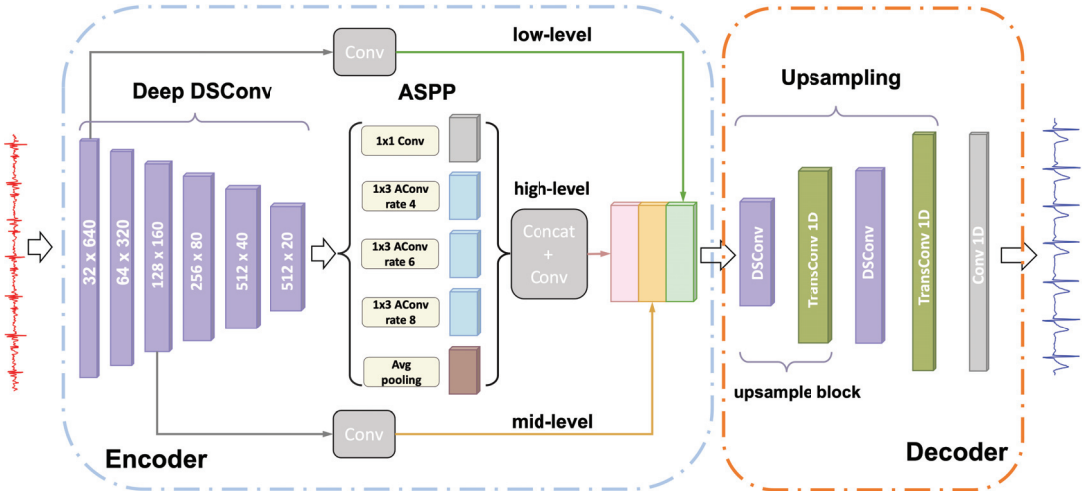


Fig. 6. Schematic of the Multi-level ECG-recovery Model. It is an encoder-decoder architecture that takes a processed SCG sequence as input and an ECG sequence as output.

Normalization method [52] to normalize the SCG magnitude measurements, which can be described as:

$$S_{scaled}(n) = \frac{S'_n - \mathrm{Min}(S'_n)}{\mathrm{Max}(S'_n) - \mathrm{Min}(S'_n)} \tag{1}$$

where S'_n denotes the magnitude of the n-th point in filtered SCG measurements after *Soft-Thresholding*. In this way, all the SCG readings will be scaled to the same level (i.e., $0 \sim 1$), indicating that the impacts of posture and pressing strength have been significantly mitigated. Note that we perform normalization to all SCG measurements before the ECG recovery.

4.2 Multi-level ECG-recovery Model

In this section, we first formally define the ECG recovery problem and notations and then detail the architecture of our proposed Multi-level ECG-recovery Model (MEM).

4.2.1 Problem Definition. In data recording, the accelerometer emits univariate SCG time series $\boldsymbol{X} = \{\boldsymbol{x}_i\}_{i=0}^{t}$, where t is the number of SCG time steps. The ECG recovery problem is then defined as:

$$\hat{Y} = f(\boldsymbol{X}), \text{ where } \hat{\boldsymbol{Y}} = \{\hat{\boldsymbol{y}}_i\}_{i=1}^{p} \tag{2}$$

is the recovered ECG signals. The length p of the recovered ECG signals is arbitrary compared to the length t of input SCG signals. Our aim is to find an optimal model f^* so that the discrepancy between the recovered ECG $\hat{\boldsymbol{Y}}$ and the actual ECG $\boldsymbol{Y}$ is minimized:

$$f^* = \arg\min_{f} \mathcal{L}(\boldsymbol{Y}, \hat{\boldsymbol{Y}}), \tag{3}$$

where $\boldsymbol{Y} = \{\boldsymbol{y}_i\}_{i=1}^{p}$ is the actual ECG signals recorded from a professional ECG instrumentation and $\mathcal{L}(.,.)$ is the loss function that measures the difference between $\boldsymbol{Y}$ and $\hat{\boldsymbol{Y}}$.

4.2.2 Recovery Model. In light of the above problem definition, the ECG recovery problem can be cast as a sequence-to-sequence transformation. Encoder-decoder architectures have achieved great success in a sequence-to-sequence transformation in various domains [53–55]. Typically, an encoder-decoder architecture is composed

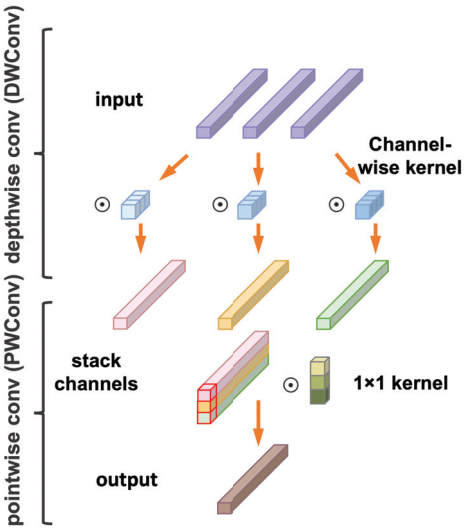


Fig. 7. Schematic of depthwise separable convolution.

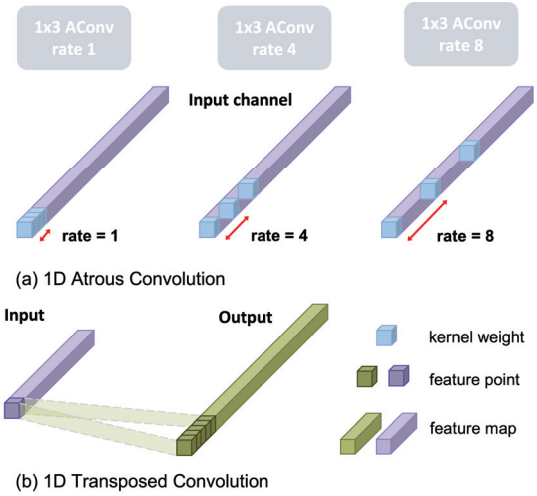


Fig. 8. Schematic of (a) 1D atrous convolution and (b) 1D transposed convolution.

of (1) an encoder module that extracts key information while filtering redundant information, and (2) a decoder module that reconstructs desired both coarse-grained and fine-grained details from the distilled information. Our MEM follows the encoder-decoder routine but with specialized encoder and decoder modules that can successfully explore the underlying vibration of heartbeats and recover them into ECG waveform.

Encoder with depthwise separable convolution and ASPP. As shown in Fig. 6, our encoder module starts with a deep *Depthwise Separable Convolution (DSConv)* module. The DSConv features a *depthwise convolution (DWConv)* followed by a *pointwise convolution (PWConv)* (see Fig. 7) to achieve reduced computational complexity while competitive or even better performance compared to conventional convolutional networks. As shown in Fig. 7, DWConv executes convolution on each input channel independently. We apply 1D DWConv to the input sequence:

$$\mathrm{DWConv}(\mathbf{W}, \boldsymbol{X})_{(i,k)} = \sum_{m=0}^{M} \mathbf{w}_{(m,\,k)} \cdot \boldsymbol{x}_{(i+m,\,k)}, \tag{4}$$

where $\mathrm{DWConv}(\mathbf{W}, \boldsymbol{X})_{(i,k)}$ is the i-th element of the k-th output channel, $\boldsymbol{X}$ is the input sequence, $\mathbf{W}$ is the weight matrix, and M is the kernel size. The PWConv is a simple 1×1 convolutional network (see Fig. 7) with minimal computation and model parameters to fuse information from each output channel of DWConv. The 1D PWConv is formalized as follows:

$$\mathrm{PWConv}(\mathbf{W}, \boldsymbol{X})_{(i,q)} = \sum_{k=0}^{K} \mathbf{w}_{(i,k)} \cdot \boldsymbol{x}_{(i,\,k)}, \tag{5}$$

where $\mathrm{PWConv}(\mathbf{W}, \boldsymbol{X})_{(i,q)}$ is the i-th element of the q-th output channel and K is the number of input channels. Then, by combining DWConv and PWConv, a DSConv block is:

$$\mathrm{DSConv}(\mathbf{W}_p, \mathbf{W}_d, \boldsymbol{X})_{(i,q)} = \mathrm{PWConv}(\mathbf{W}_p, \mathrm{DWConv}(\mathbf{W}_d, \boldsymbol{X})_{(i,k)})_{(i,q)}, \tag{6}$$

where $\mathbf{W}_d$ and $\mathbf{W}_p$ are weight matrices for DWConv and PWConv, respectively. Our deep DSConv module has six cascade DSConv blocks with a progressively doubled number of channels {32, 64, 128, 256, 512, 512} and halved

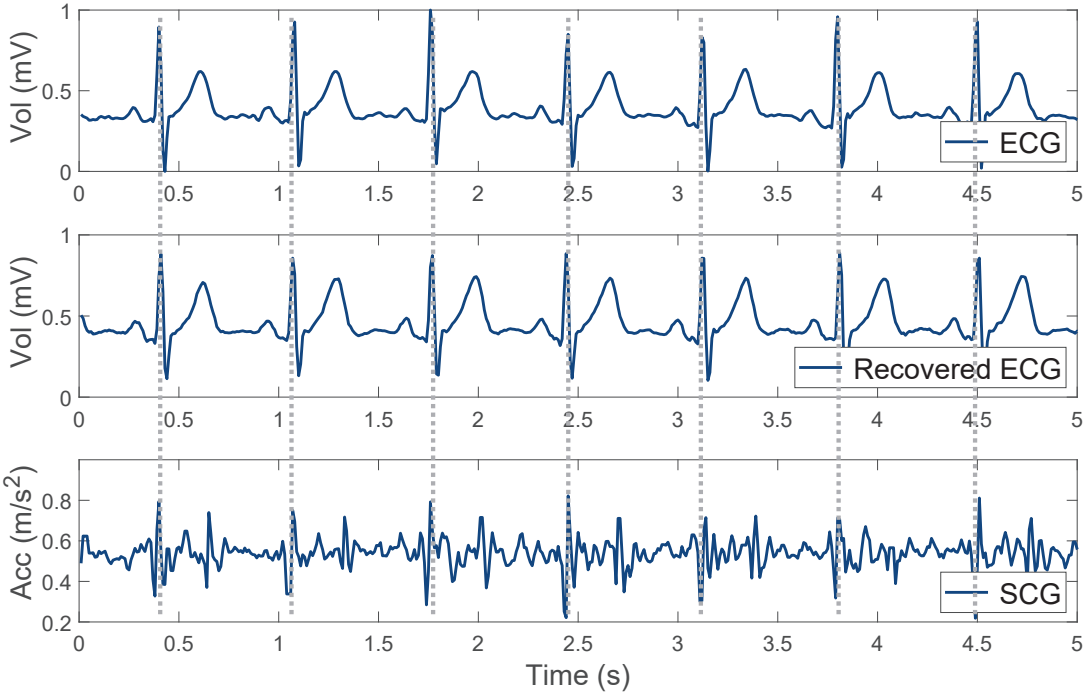


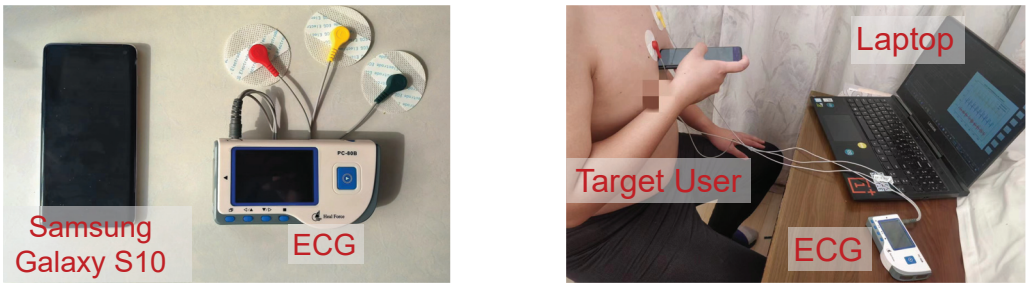
Fig. 9. Comparison between ground truth and recovered ECG waveform.

channel size {640, 320, 160, 80, 40, 20}. After each DSConv block, we insert a Batch Normalization layer [56] and Leaky ReLU activation [57]. This deep structure is able to extract features with gradually refined granularity. To fully recover both the coarse-grained fluctuations and fine-grained details of ECG, we take features of three sorts of granularity, low-level features, mid-level features, and high-level features, as input into the subsequent decoder (see Fig. 6). For each feature level, we use a convolution layer to cast the feature size to be compatible for later concatenation.

Since the high-level feature contains the most condensed and fine-grained information, we employ an Atrous Spatial Pyramid Pooling (ASPP) module [55, 58] for a more aggressive representation. As shown in Fig. 6, the ASPP module employs multiple parallel 1D atrous convolution (AConv) with pyramid atrous rates to explore multi-scale representations. It also has a PWConv for the finest-grained representation and a global average pooling for the coarsest-grained representation. The AConv enables magnifying the receptive field of filters with normal computation and parameter amounts:

$$\mathrm{AConv}(\mathbf{W}, \boldsymbol{X})_{(i,q)} = \sum_{k=0}^{K} \sum_{m=0}^{M} \mathbf{w}_{(m,\,k)} \cdot \boldsymbol{x}_{(i+r\cdot m,\,k)}, \tag{7}$$

where $\mathrm{AConv}(\mathbf{W}, \boldsymbol{X})_{(i,q)}$ is i-th element of the q-th output channel, r is the atrous rate, M is the kernel size, and K is the number of input channels. Fig. 8(a) depicts the schematic of atrous convolution with various atrous rates. An atrous convolution with an r atrous rate has $r-1$ holes (i.e., without kernel values) between successive kernel values. This setting is able to expand the receptive field of conventional convolution from $k \times k$ to $k + (r-1)(k-1)$ without extra network parameters and computation. In this way, the ASPP module offers an extraordinary competence to deal with both short- and long-range dynamics of input sequences. We carefully select the atrous rate as 4, 6, and 8



(a) Devices (b) Scenario

Fig. 10. Experimental setup.

considering the high-level feature size. DSConv and AConv in the encoder enable a lightweight ECG recovery model which is desired by mobile devices.

Decoder. The decoder takes the concatenated multi-level features as input and upsamples it to the size of the corresponding actual ECG signals. As shown in Fig. 6, it consists of two *upsample blocks*, each of which has one DSConv layer for feature extraction and one 1D Transposed Convolution (*TransConv 1D*) layer for upsampling. Through the decoding process, the size and number of features are gradually enlarged and reduced, respectively. Different from naive interpolation technology, TransConv upsamples the input feature map to a desired size of the output feature map using learnable filters. Fig. 8 (b) shows the schematic of a 1D TransConv layer. We take a 1×5 kernel as an example (Fig. 8 (b)): one input feature point will multiply with every weight of the kernel, resulting in a 1×5 output feature. By combining a DSConv and TransConv, an upsample block can well interpret the condensed features from the encoder and recover them into desired ECG signals.

4.2.3 Training and Validation Scheme. We use the Huber loss [59], which makes the best of both the mean squared error (MSE) and the mean absolute error (MAE), as the training loss function in Eq. 3. It is defined as:

$$\mathcal{L}(\boldsymbol{y}_i, \hat{\boldsymbol{y}}_i) = \begin{cases} 0.5(\boldsymbol{y}_i - \hat{\boldsymbol{y}}_i)^2 & \text{if } |\boldsymbol{y}_i - \hat{\boldsymbol{y}}_i| < 1, \\ |\boldsymbol{y}_i - \hat{\boldsymbol{y}}_i| - 0.5 & \text{otherwise.} \end{cases} \tag{8}$$

The Huber loss has shown extraordinary robustness on various regression tasks [60–62]. It penalizes linearly for large errors while quadratically for small errors so that can accelerate the convergence speed and ignore noise at the beginning of training and then gradually reduce optimization steps to ensure global minima. We use the Adam optimizer for training with possible early stopping. After full training, we obtain an end-to-end model with SCG as input and recovered ECG as output. Due to factors like smartphone positions (subject-agnostic) and body weight (subject-related), the correspondence between SCG and ECG varies from one subject to another making it challenging to build a unified recovery model for different subjects. However, it is always preferred to directly apply the trained model without lengthy subject-specific calibration or model retraining for practical use. Therefore, different from Cao et al. [28], we validate our ECG recovery system in a subject-independent setting, where training and test are performed on different subjects. Our experimental results, as detailed in Section 6, well justify the effectiveness of our recovery system. There are two reasons accounting for our outstanding subject-independent performance. First, we construct a large dataset consisting of many subjects for model training, which lays the foundation for our ECG recovery model to extract the common patterns of SCG-to-ECG correspondence across different subjects. Second, we propose a Multi-level ECG-recovery Model, which is capable of extracting low-, mid-, and high-level patterns of SCG-to-ECG correspondence. Unlike extracting only high-level deep features and

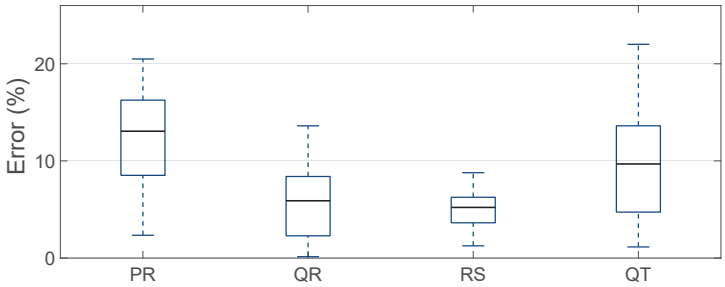


Fig. 11. Estimation error of amplitude in ECG recovery.

thus being prone to overfitting specific subjects, our multi-level structure allows for better extraction of common patterns across different subjects and facilitates direct application to unseen subjects. In addition, the ASPP module is capable of capturing not only short-term dynamics but also gradually increasing long-term dynamics, enhancing feature extraction capabilities. By training it with the subject-diverse dataset, our Multi-level ECG-recovery Model can well discover the underlying subject-independent patterns of SCG-to-ECG correspondence.

5 IMPLEMENTATION

We have implemented our system on a Samsung Galaxy S10 mobile phone with a built-in accelerometer and running Android 10.0 OS, as shown in Fig. 10(a). To facilitate evaluation, we use the mobile phone to collect data, which is then transferred to a PC for further processing. The sampling rate is set to 100 Hz, which is sufficient to cover the frequency range of heartbeat activity. We utilize a commercial 3-lead ECG monitor (Heal Force PC-80B), which has proven accurate [63] and been extensively employed in research [64–66], to obtain the ground truth [67], which serves as the reference for experimental evaluation and as labels for training the ECG recovery model. Given that the ECG device's default sampling rate is 150 Hz, we employ cubic spline interpolation to decrease the sampling rate to 100 Hz, which matches the sampling rate of the mobile phone. The accelerometer data are then sent to a LENOVO Y7000P PC via WiFi for further processing in Python 3.9, including denoising, data augmentation, deep learning model training, and ECG recovery.

The proposed MEM is trained on NVIDIA GeForce GTX 1060 GPU with 6 GB of GDDR5 memory. Next, we report specific parameter settings of the MEM. The kernel size and the stride size of DWConv in the Deep DSConv module are 1×3 and 2, respectively, to gradually shrink the feature size. The transposed convolution layer uses a kernel size of 2, which expands the input feature size by a factor of 2. We use the batch normalization and Leaky ReLu activation function in the Deep DSConv module. Dropout with a rate of 0.5 is used to alleviate potential overfitting. The model is trained using the Adam optimizer with an annealing learning rate starting from 10^{-3} and ending at 10^{-6}. The training proceeds for 500 epochs with possible early stopping.

6 EVALUATIONS

6.1 Experimental Setup

To evaluate system performance, we recruit 30 participants (15 females and 15 males) aged between 22 to 83 years. The average age of the participants is 45.9 years, with a standard deviation of 22.18. Of the 30 participants, 15 are healthy, and the remaining 15 have been diagnosed with a range of heart conditions including tachycardia, bradycardia, arrhythmia, unstable angina, and myocardial infarction. Each disease is represented by three participants. All participants have not participated in similar experiments before. Their heart rate is ranged from 53 to 125 BPM (Beats Per Minute) with an average of 85 BPM. We start each experiment by providing a

Table 1. Overall performance of the MEM (ours) and comparison with the baselines. The best results are shown in bold.

Metric	Model	Sub 1	Sub 2	Sub 3	Sub 4	Sub 5	Sub 6	Sub 7	Sub 8	Sub9	Sub 10	Mean
CS	AE	0.9386	0.9276	0.8148	0.806	0.8098	0.8381	0.861	0.9446	0.9344	0.8466	0.8722
	VAE	0.8989	0.9101	0.8068	0.7800	0.8987	0.7480	0.8648	0.955	0.9138	0.7204	0.8497
	LSTM	0.9093	0.9015	0.5904	0.7984	0.9398	0.824	0.8503	0.9511	0.9172	0.8712	0.8553
	GRU	0.9104	0.9046	0.6049	0.8053	0.9812	0.8517	0.8414	0.9525	0.9537	0.9149	0.8721
	Seq2Seq	0.9247	0.9154	0.8201	0.8222	0.9071	0.8301	0.9112	0.9362	0.9404	0.9026	0.8910
	Ours	**0.9455**	**0.9351**	**0.8302**	**0.8543**	**0.9582**	**0.8514**	**0.9598**	**0.9800**	**0.9634**	**0.9187**	**0.9197**
PR (ms)	AE	17.16	15.30	25.61	19.75	7.182	11.97	18.99	12.33	27.43	7.488	16.32
	VAE	36.42	35.72	37.67	26.83	38.27	44.51	40.01	37.26	46.43	40.47	38.36
	LSTM	45.10	55.64	51.11	39.76	36.27	29.99	44.57	43.75	48.16	26.64	42.10
	GRU	41.37	49.38	50.79	40.63	33.71	26.63	40.92	37.51	43.07	22.73	38.67
	Seq2Seq	10.62	19.37	31.22	22.53	9.730	12.11	25.09	11.43	27.62	7.911	17.76
	Ours	**8.997**	**11.79**	**17.12**	**15.98**	**5.413**	**8.596**	**16.67**	**10.34**	**19.01**	**6.301**	**12.02**
QRS (ms)	AE	19.232	17.34	45.92	18.13	31.65	29.94	26.04	17.77	51.53	26.67	28.42
	VAE	117.2	143.7	121.2	85.97	131.3	96.56	111.9	97.81	93.25	106.5	110.5
	LSTM	118.3	141.9	90.9	140.8	143.4	113.1	114.1	119.7	168.7	85.21	123.6
	GRU	164.7	131.6	57.06	124.2	89.72	90.12	135.7	53.76	173.8	100.7	112.1
	Seq2Seq	28.78	25.64	42.01	20.65	36.89	27.21	30.64	25.67	38.32	23.94	29.98
	Ours	**14.221**	**16.25**	**16.98**	**21.28**	**17.50**	**16.19**	**14.64**	**16.69**	**18.58**	**16.71**	**16.90**
QT (ms)	AE	30.62	21.3	33.74	23.76	21.75	21.82	32.25	33.13	25.81	31.8	27.60
	VAE	52.41	65.55	61.74	65.51	61.81	91.65	81.01	65.91	91.62	86.31	72.35
	LSTM	105.1	97.07	59.48	67.05	81.93	85.18	63.71	107.5	127.83	63.77	85.86
	GRU	90.88	82.34	71.13	60.49	86.47	76.46	81.44	86.64	94.71	70.95	80.15
	Seq2Seq	42.03	31.22	24.72	25.94	31.63	17.25	22.04	27.62	20.16	25.51	26.81
	Ours	**20.03**	**19.71**	**14.24**	**19.73**	**11.39**	**15.81**	**20.52**	**15.73**	**14.32**	**14.9**	**16.64**
RR (ms)	AE	1.987	4.096	6.633	0.485	0.765	4.092	1.503	2.803	4.807	2.444	2.962
	VAE	37.41	49.34	21.72	17.44	15.92	26.73	15.82	41.63	40.64	37.02	30.37
	LSTM	15.27	18.32	14.54	21.26	17.92	32.58	10.80	23.70	19.33	13.91	18.76
	GRU	19.72	23.61	12.08	5.428	14.51	22.28	4.326	3.918	12.43	9.872	12.82
	Seq2Seq	5.931	11.78	10.58	1.211	15.19	7.514	2.229	5.933	3.434	6.829	7.063
	Ours	**0.733**	**3.281**	**3.064**	**0.919**	**0.536**	**2.845**	**0.687**	**1.127**	**2.874**	**2.337**	**1.840**

10-minute briefing session to introduce the experimental equipment and data collection methods. Ethical approval has been obtained from our institutional review board (IRB) prior to all experiments. For each subject, data collection lasts for 2 minutes and is repeated 10 times. If not specified, each subject is instructed to remain still and use a smartphone to capture the SCG signal by placing it against their chest, as shown in Fig. 10(b). Unless otherwise specified, each subject participates in all experiments, and we use leave-one-out cross-validation to evaluate system performance with a user-independent data set. Specifically, for n subjects, we divide the raw data into n sets, and the testing is iterated n times. In each iteration, we use the set of $(n-1)$ subjects as the training samples and the set of the remaining subject as the testing samples. Unless otherwise specified, for all experiments, we use the QRS interval, a commonly used indicator in medical diagnosis [68, 69], and cosine similarity, which is to measure the morphological similarity (a common function in MATLAB) between the recovered ECG signal and the ground truth, to evaluate the system's performance.

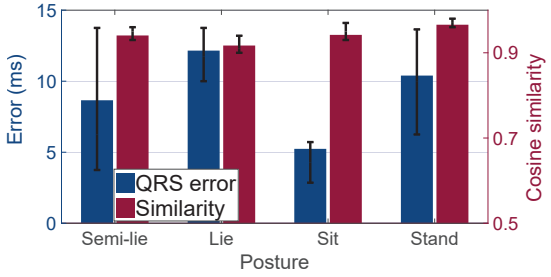


Fig. 12. Error for different postures.

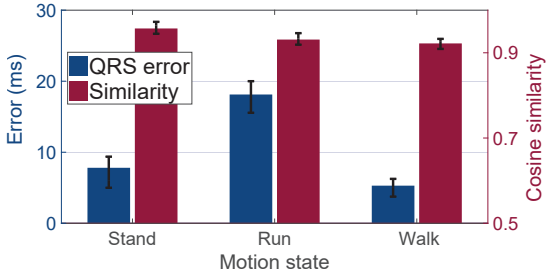


Fig. 13. Error after different movements.

6.2 Overall Performance

To evaluate system performance, we define five evaluation metrics. The first one is a coarse-grained recovery metric, cosine similarity (CS), which measures the similarity between the recovered and ground truth signals, with a higher value indicating better performance. The remaining four metrics are fine-grained clinical recovery metrics, which evaluate the difference between the recovered and ground truth values of key ECG intervals including PR interval, QRS interval, QT interval, and RR interval. Lower values of these metrics indicate better performance. These intervals correspond to specific distances between the peaks and valleys in the ECG waveform, and they are important diagnostic references for various heart diseases[25].

Table 1 shows the overall performance of our system (ours) and its comparison to using alternative popular deep learning models for sequence-to-sequence transformation, autoencoders (AE), variational autoencoders (VAE), Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and Sequence-to-Sequence (Seq2Seq). Note that for some cardiac diseases, it is impractical to measure some key ECG intervals so physicians diagnose them by observing the ECG wave pattern. For example, the key diagnostic feature of unstable angina is an inverted T wave (see Fig. 24) so the QT interval cannot be achieved. Therefore, we only include healthy subjects in the overall performance evaluation but present the ECG recovery evaluation for patients in Section 6.6. In addition, due to the page limit, we present the evaluation results of 10 randomly selected out of 15 healthy subjects in Table 1. Our system achieves a high resolution of ECG recovery, around 10 ms, compared to the normal PR/QRS/QT/RR intervals which have resolutions in tens to hundreds of milliseconds. This demonstrates the outstanding capability of our system in fine-grained ECG recovery, making it suitable for real-world heart disease diagnosis applications. Table 1 clearly illustrates that while our system and its comparison models have similar performance in coarse-grained recovery as measured by CS, our proposed system outperforms them in all fine-grained evaluation metrics. For example, the Seq2Seq has a similar coarse-grained recovery performance to our system (0.8910 vs 0.9197) but significantly worse fine-grained recovery performance (PR: 17.76 vs 12.02; QRS: 29.98 vs 16.90; QT: 26.81 vs 16.64; RR: 7.063 vs 1.840). This highlights the importance of designing specific ECG recovery models, such as our MEM, rather than simply using existing models from other domains. Additionally, it is important to note that the acceptable measurement deviation for QRS interval, which is defined as the time interval from Q wave to S wave, in clinical practice is 10-20 ms [70–72], and our average measurement error is 16.90 ms, which perfectly meets this accuracy requirement. While this deviation could potentially result in misdiagnosis of certain heart conditions, such as bundle branch block [27, 73, 74] (given that the QRS interval of an incomplete right or left bundle branch block may be approximately 20 ms wider than a normal QRS interval), physicians take into account of other factors such as ECG waveforms (which our system recovers exceptionally well with a cosine similarity of 0.9197) and the patient's clinical presentation during the diagnostic process. As a result, even with some degree of measurement error, a reasonably accurate diagnosis can typically be achieved.

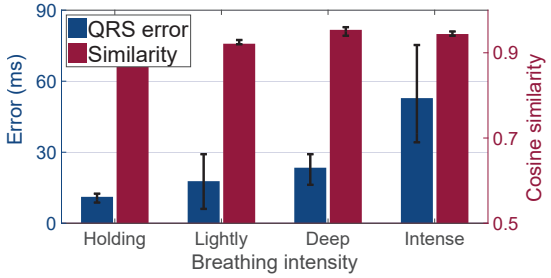


Fig. 14. Robustness over different breathing intensities.

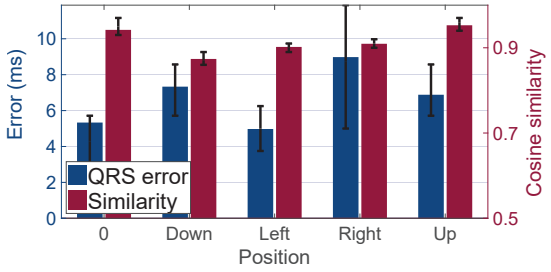


Fig. 15. Robustness over different positions for data collection.

In addition to the time interval, we also evaluate amplitude changes of the key intervals in the ECG, such as PR, QR, RS, and QT intervals. Specifically, we compute the relative percentage change in amplitude during these intervals for both recovered and ground truth ECG signals, and subsequently calculate the amplitude estimation error. As illustrated in Fig. 11, the average amplitude errors during PR, QR, RS, and QT intervals are 13.06%, 5.9%, 5.21%, and 9.69%, respectively. The average error across these intervals is 8.36%, which falls within the acceptable range for amplitude estimation according to previous studies [75].

6.3 Evaluation on Usage Scenario

In this section, we investigate the performance of our system under different application scenarios, including various postures and motion states. Since it is challenging for the system to directly remove noise interference caused by body movements, all experiments are conducted with the body keeping still during data collection.

Posture: In order to evaluate the impact of body posture, we collect data while the subjects keep sitting, standing, semi-lying, and lying down. Data collected while sitting serves as the standard training set to train the system. As shown in Fig. 12, we observe that the states of standing and semi-lying have minimal interference with heart rate recovery. However, in the semi-reclining state, the system's average measurement error is slightly increased from 5.28 ms to 8.66 ms compared to the sitting state. This is reasonable since the phone's level is not horizontal with respect to the ground, and the signal-to-noise ratio (SNR) along the Y-axis decreases, leading to an increase in the influence of hand movement. Additionally, the recovered heartbeat waveform in all three postures is highly similar to the actual waveform, showing that the system can work well in different postures.

Motion: Due to the extreme sensitivity of the built-in accelerometer, we can only evaluate the robustness of the system when the user is relatively stationary. Here, we are just committed to verifying the system's recovery capability after different motion states. We first have each user maintain three different motion states, namely standing, walking, and running, and collect their data after 1 minutes. Note that we use the samples in a all motion state as the training dataset to involve the data diversity. As illustrated in Fig. 13, we observe that the variations in heart rate following walking do not influence the recovery of the QRS interval or the waveform similarity. However, the recovery performance after running declines to 18.13 ms. This decline is predominantly due to the body remaining in a heightened state for a minute following the run, causing the accelerometer signal to be more susceptible to bodily movement. Nevertheless, even with this running-induced error, the accuracy is sufficient for everyday monitoring needs. These results suggest that our system delivers robust ECG recovery performance across varying motion states.

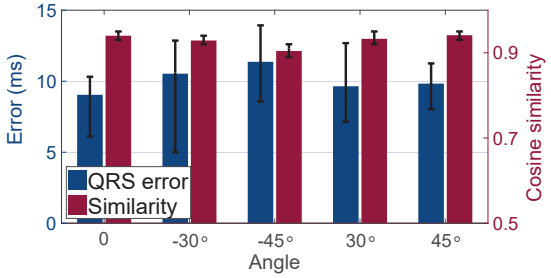


Fig. 16. Robustness over different angles for data collection.

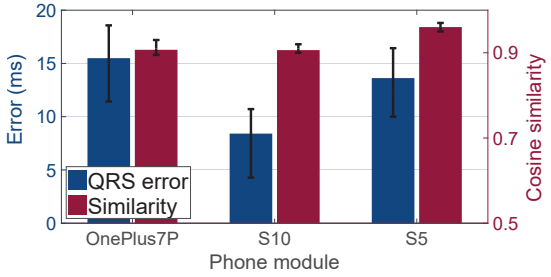


Fig. 17. Robustness over different phone modules.

6.4 Evaluation on Robustness

In this section, we assess the system's robustness in recovering ECG waveforms under various conditions. We use the original standardized data as the training data and newly collected data under different conditions as the testing data. For example, when evaluating the system's robustness to angle changes, we use data collected by mobile phones at different angles as test data.

Respiration: To investigate the impact of respiration on the system, we direct the subjects to breathe at different levels, namely, holding breath, normal breathing, deep breathing, and intense breathing. We consider the data obtained from normal breathing as the standard training sample. As illustrated in Fig. 14, the average QRS interval errors for holding breath, normal breathing and deep breathing are 11.16 ms, 17.78 ms, and 23.75 ms, respectively, which are sufficient for daily monitoring. However, the error jumps to an unacceptable 52.85 ms during intense breathing, while the similarity between the recovered waveform and the ground truth displays a corresponding decreasing trend. This can be mainly attributed to the fact that intense breathing is typically accompanied by body movements. In general, our system is capable of withstanding the effects of normal and deep breathing.

Position: To evaluate the impact of pressing position on ECG recovery, we first mark the reference point's position on the volunteer's clothing. Next, we instruct the volunteers to press the phone 5 cm to the left, right, up, and down from the reference point to measure the QRS interval errors. We use the samples collected from the reference point as training data and the samples collected from the other positions as testing data. As shown in Fig. 15, the error in the left position is similar to that of the reference point. This is reasonable since the heart is located to the left of the chest. However, the performance of the other three positions has decreased to varying degrees, mainly due to a reduction in amplitude in the SCG signals. It is worth noting that although the measurement performance has declined, it is still sufficient for daily monitoring needs.

Angle: Under standard conditions, we define the angle when the user holds their phone vertically against their chest as 0°. When the phone is facing upward, the angle is defined as positive and negative when facing downward. As shown in Fig. 16, we observe that when the angle is in the range of 0° to 45°, our system achieves equally excellent performance in recovering the heartbeat. This indicates that the angle has little effect on heartbeat recovery in the range of $0^\circ \sim 45^\circ$. However, when facing downward, the impact slightly increases, resulting in a larger error. This is mainly due to the fact that the heartbeat is mostly concentrated above. The upward phone can collect signals almost identical to those collected in normal conditions. However, sufficient quality signals cannot be obtained in the downward direction, leading to a low SNR. Overall, our system can tolerate angle changes in the range of 75° (i.e., $-30^\circ \sim 45^\circ$). Therefore, when the user slightly adjusts the phone's angle, the collected heartbeat samples can still be used for recovery.

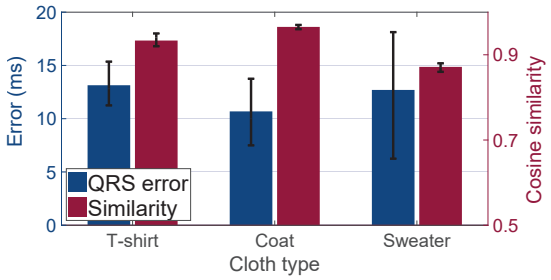


Fig. 18. Robustness over different types of clothes.

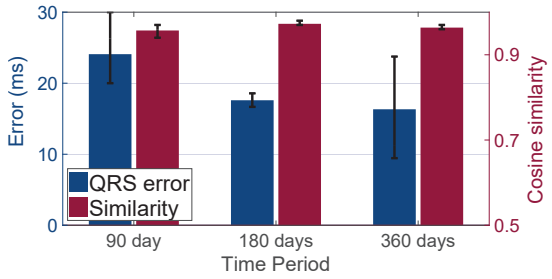


Fig. 19. Robustness over time.

Phone type: We deploy our system on three smartphones, Samsung Galaxy S5, Samsung Galaxy S10, and OnePlus7P, to study the recovery capabilities of different phone modules. As shown in Fig. 17, the average errors of the QRS interval for the Samsung Galaxy S5, Samsung Galaxy S10, and OnePlus7P are 13.63 ms, 8.42 ms, and 15.51 ms, respectively. Despite exhibiting significantly more noise than the Samsung Galaxy S10 and OnePlus 7P, the Samsung Galaxy S5 demonstrates a recovery performance that is on par with two other phone models. This is mainly due to our two-tiered noise removal scheme which can effectively remove system noise. Overall, the experimental results demonstrate the effectiveness of our system for different smartphone modules.

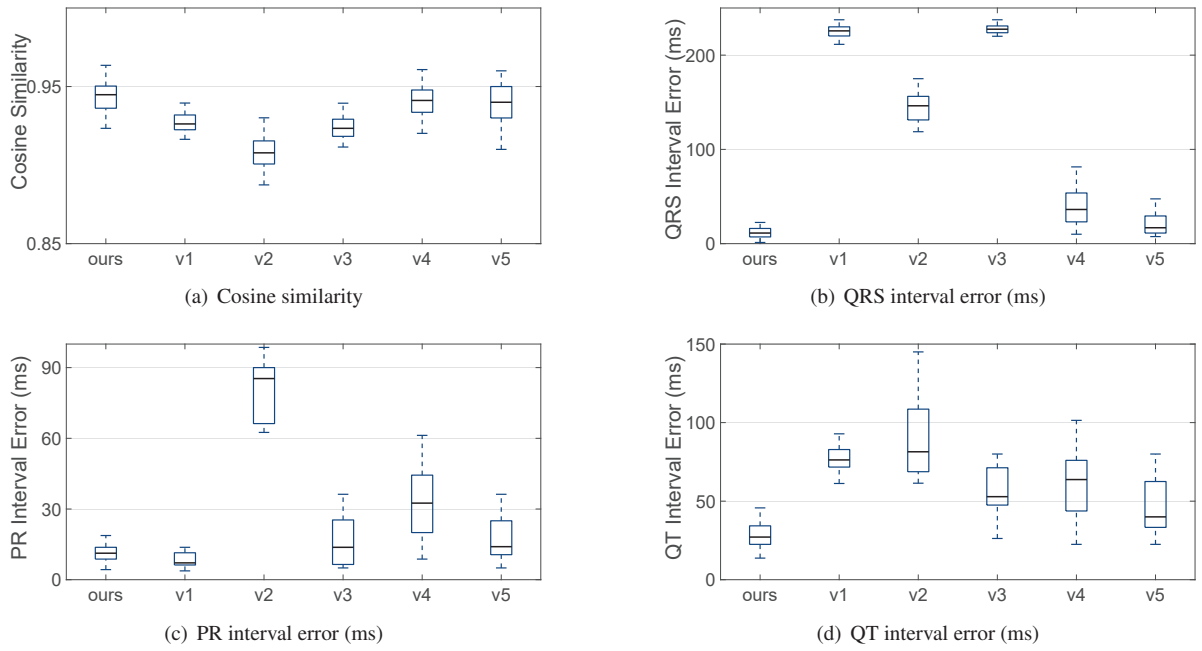
Clothing: To simulate summer, autumn, and winter clothing, we ask subjects to wear T-shirts, coats, and sweaters when collecting the accelerometer data. We use the data collected while wearing T-shirts as the standard training data. Fig. 18 illustrates that the average interval errors of the QRS complex for T-shirts, coats, and sweaters are 10.69 ms, 13.14 ms, and 12.87 ms, respectively, which is consistent with the trend of cosine similarity. Although the measurement error increases with the thickness of the clothing, our system still has the capability in recovering the ECG waveform accurately.

Time period: To explore the long-period stability of the system, we initially select a reference time and use the data collected at that time as the standard training data. We then ask the same subjects to collect data again for testing at 90 and 180 days later. Fig. 19 shows that the performance of our system is highly stable over a six-month period. This is mainly because we develop a multi-level deep learning recovery model that can effectively address the user-independent issue, which indicates that even if there is a slight change in the heartbeat pattern, the system still has strong capability in recovery.

6.5 Ablation Study

We conduct an ablation study to justify the effectiveness of each component in our deep learning recovery model. Specifically, we implement a group of variants by removing several of the components and observed their impact on all evaluation metrics. The variants include **v1** (w/o Deep DSConv) which only uses the ASPP module in the encoder, **v2** (w/o ASPP) which does not use the ASPP module, **v3** (w/o mid- and low-level features) which only has the Deep DSConv module directly connected to the decoder, **v4** (w/o mid-level features) which only has low- and high-level features, and **v5** (w/o low-level features) which only has mid- and high-level features.

Fig. 20 shows the results of the ablation study with different evaluation metrics. Note that all variants have very small RR interval recovery errors so we only show the results with the other evaluation metrics. Ours, which refers to our recovery model, generally outperforms all variants demonstrating the effectiveness of each adopted component. Specifically, ours not only has the highest cosine similarity and the lowest QRS/PR/QT interval errors but also the smallest variance regarding these evaluation metrics. This demonstrates that all components greatly contribute to the results regarding both coarse- and fine-grained ECG recovery. When comparing with **v1** and **v2**, we conclude that both the Deep DSConv and ASPP modules are essential to compelling high-level feature



(a) Cosine similarity

(b) QRS interval error (ms)

(c) PR interval error (ms)

(d) QT interval error (ms)

Fig. 20. Ablation study on different evaluation metrics. Please refer to section 6.5 for the details of each variant $\mathbf{v}*$.

extraction. The comparison with **v2**, **v4** and **v5** demonstrates the necessity of feature extraction of each level for the success of ECG recovery. Moreover, **v5** achieves the performance closest to ours indicating that the low-level features are less powerful than the other counterparts.

6.6 Case Study on Cardiac Diseases

To further evaluate the performance of our system in recovering ECG, we recruit 15 patients, each with a different heart condition, including tachycardia, bradycardia, sinus arrhythmia (SA), unstable angina (UA) and myocardial infarction (MI). Three patients are selected for each condition. The aim is to explore the feasibility of our system for early disease screening. We employ a leave-one-out cross-validation strategy to assess system performance using a user-independent dataset, as detailed in Section 6.1. For instance, to evaluate the system's performance on tachycardia, we repeat the testing process three times, corresponding to the three patients diagnosed with this condition. In each iteration, only one patient's data is used as the testing set. Ultimately, all the results are collated to evaluate the overall system performance.

Tachycardia: Tachycardia is a condition characterized by a heart rate exceeding 100 BPM, which is beyond the normal range (60 to 100 BPM) of adults at rest. It should be noted that all patients participating in the experiment are instructed to sit quietly and maintain emotionally stable to mitigate potential physiological effects. Fig. 21 presents a comparison between the ground truth and the recovered ECG waveforms, which show the heart rates of 104.8 BPM and 104.3 BPM, respectively, as calculated using the RR interval. The high degree of similarity between the recovered waveform and the ground truth is evident. As demonstrated in Figure 26, the average QRS interval error for tachycardia is 24.65 ms, and the average similarity is 95.29%. These results show that the system is capable of accurately recovering ECG signals for patients with tachycardia.

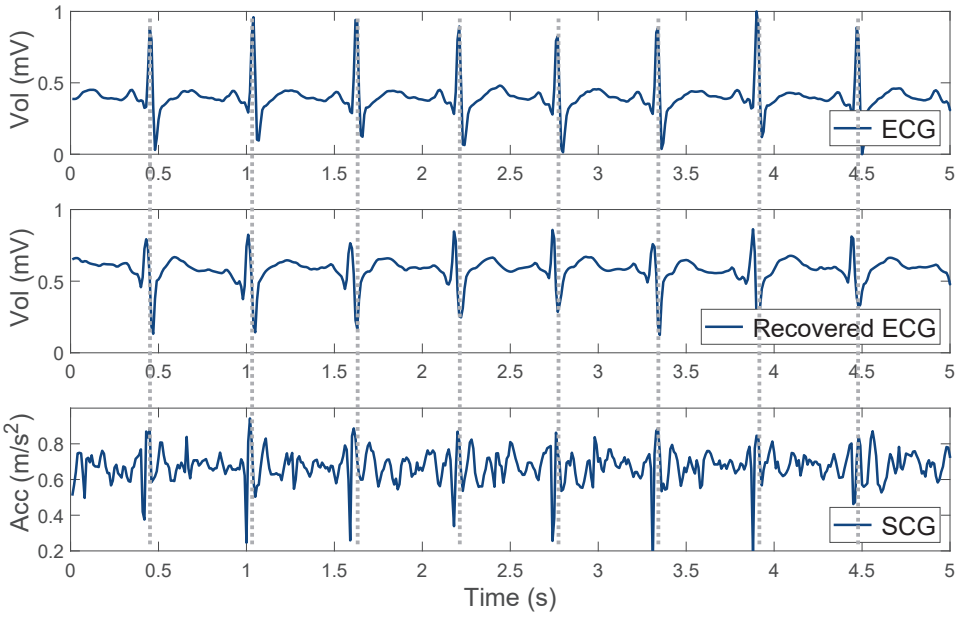


Fig. 21. Recovered ECG waveform from a patient with tachycardia.

Bradycardia: In contrast to tachycardia, bradycardia is diagnosed when an adult's resting heart rate falls below 60 BPM. It is essential to diagnose bradycardia promptly and accurately since tachycardia can potentially develop into conditions such as pathological sinus bradycardia, sinus arrest, sinus block, and atrioventricular block. Fig. 22 compares the waveforms between the ground truth and the recovered ECG with heart rates of 53.8 BPM and 53.7 BPM, respectively. The similarity in their heartbeat patterns is evident. As shown in Fig. 26, the average QRS interval error for bradycardia is 19.01 ms with an average similarity of 97.19%.

Sinus arrhythmia: Sinus arrhythmia, a cardiac condition characterized by irregularities in heartbeat rhythm, can potentially lead to serious complications such as myocardial infarction, heart failure, or stroke if not accurately and timely diagnosed. The diagnosis of sinus arrhythmia is based on the intervals between waveforms in the ECG signal (i.e., RR interval). As depicted in Fig. 23, our system can successfully recover ECG waveform even when the raw SCG data has no obvious patterns. The experimental results in Fig. 26 demonstrate an average QRS interval error of 13.87 ms and a similarity of 86.53%. Although there is a slight reduction in similarity, the accuracy is adequate for preliminary arrhythmia diagnosis.

Unstable angina: Unstable angina is a cardiac condition that poses a significant risk of progressing rapidly into a myocardial infarction or heart attack, leading to irreversible heart damage and potential fatality. Thus, it is essential to accurately and timely diagnose unstable angina. In the ECG waveform, unstable angina is usually recognized by the T wave inversion. As illustrated in Fig. 24, our system demonstrates excellent capabilities in accurately recovering the ECG waveform for unstable angina, including the distinctive T-wave inversion. These results suggest the potential of our system for the preliminary diagnosis of unstable angina. As shown in Fig. 26, the average QRS interval error and similarity in estimating unstable angina are 25.91 ms and 96.03%, respectively.

Myocardial infarction: Myocardial infarction, a widespread cardiac condition, poses a significant threat as it may result in permanent heart damage, potentially leading to complications such as heart failure, or even sudden death. In the ECG, myocardial infarction is typically characterized by T-wave inversion, similar to unstable angina.

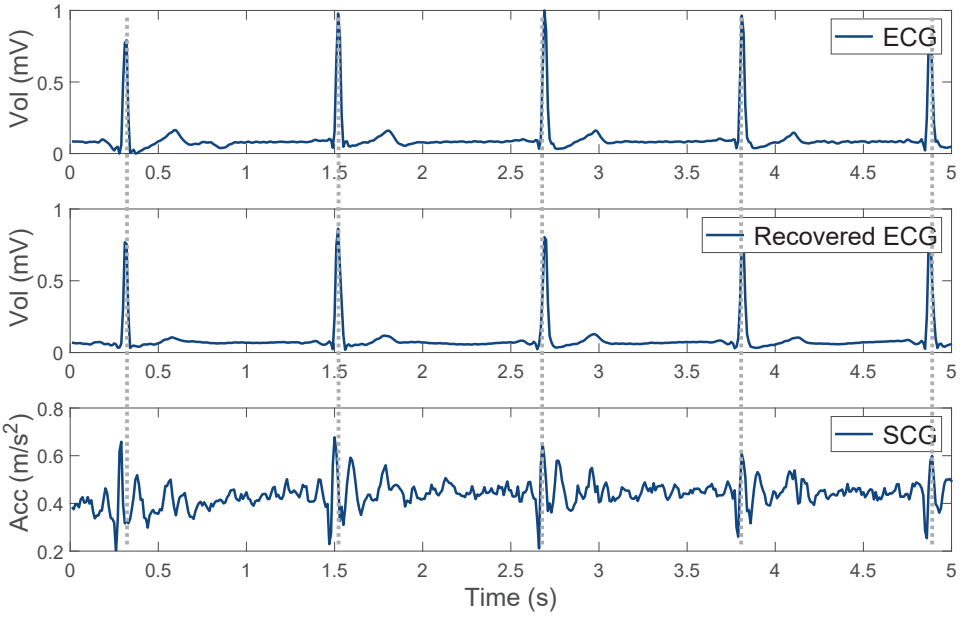


Fig. 22. Recovered ECG waveform from a patient with bradycardia.

However, a distinguishing feature of myocardial infarction is the presence of ST-segment elevation [76]. As shown in Fig. 24, a subtle elevation is noticeable during the ST-segment in the ground truth ECG signal. Impressively, our system effectively recovers the distinctive heartbeat pattern, including both T-wave inversion and ST-segment elevation. The robust performance of our system is further demonstrated by the statistical results presented in Fig. 26, which indicate an average QRS interval error of 22.81 ms and an average similarity of 93.12%, respectively.

6.7 Unseen Heart Conditions

Although the disease tests previously were user-independent, it required the model to contain samples of the same disease during training. In this subsection, we study the system's ability to recover ECGs for diseases it has never seen before. We still use the leave-one-out method, that is, when testing tachycardia cases, all of these cases are used for testing. As shown in the Fig. 27, we found that the system's recovery capability for tachycardia, bradycardia, and sinus arrhythmia is competitive with previous results. However, the recovery capability for Unstable angina and myocardial infarction significantly decreases. This is mainly because the first three diseases primarily involve changes in the occurrence cycle, with little waveform variation. In contrast, the latter two primarily involve waveform changes, and our deep learning model's ability to recover ECG for such diseases is inferior.

7 SYSTEM LATENCY AND POWER CONSUMPTION

Our system achieves low latency, making it suitable for real-time monitoring. To analyze its performance, we implement our system on a Samsung Galaxy S10 smartphone, equipped with a Qualcomm SM8150 Snapdragon 2.9 GHz Octa-core CPU. Our implementation operates on a two-thread architecture consisting of the display thread and the recovery thread. To process SCG signals, our system uses a buffer size of 500 data samples, representing a time duration of 5 seconds under a 100 Hz sampling rate). This buffer is updated every 10 ms. On average, our system takes 2.31 ms and 1.29 ms for SWT-based denoising and Soft-Thresholding, respectively, to process one

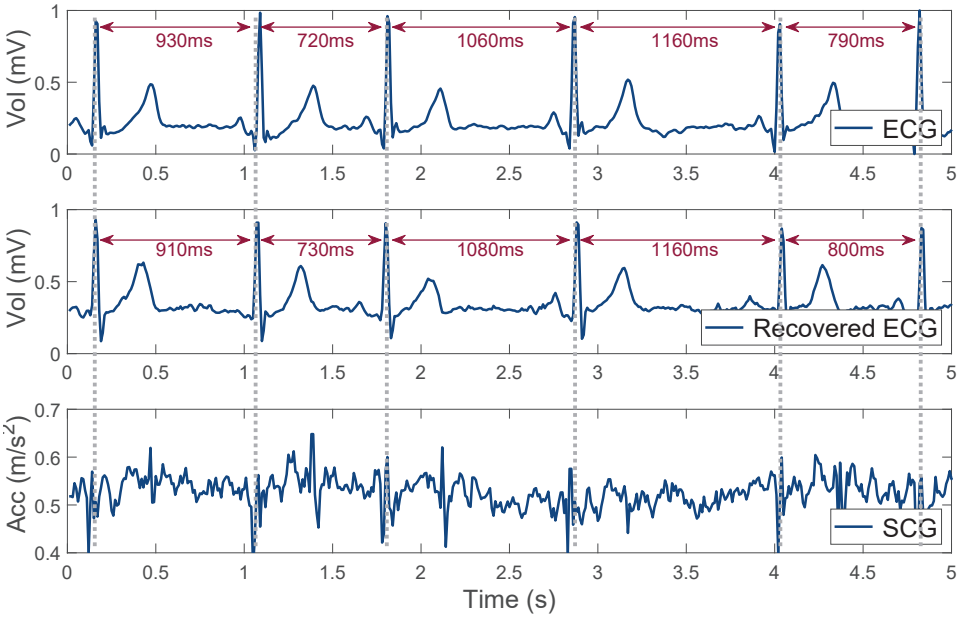


Fig. 23. Recovered ECG waveform from a patient with sinus arrhythmia.

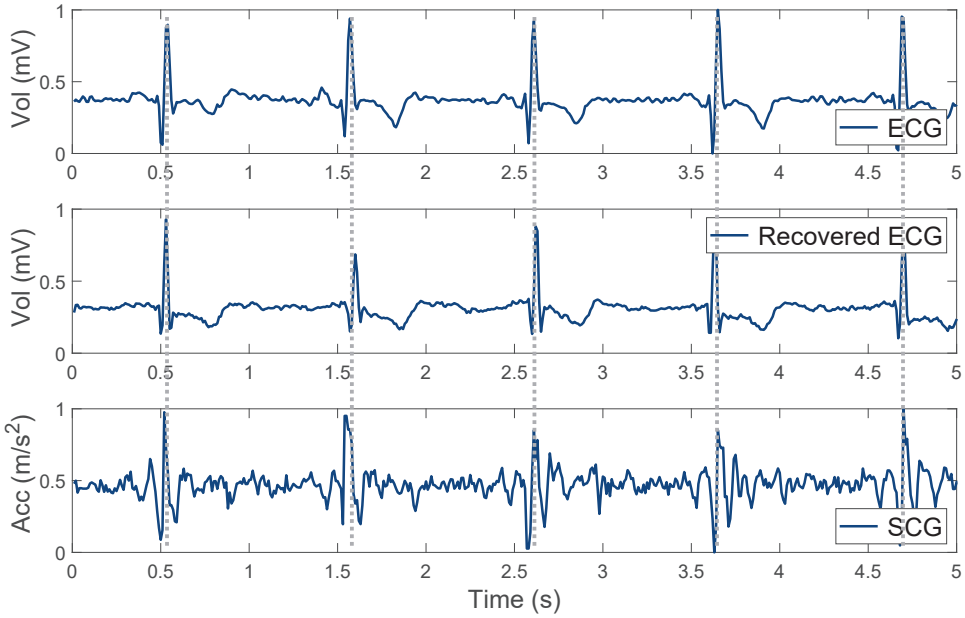


Fig. 24. Recovered ECG waveform from a patient with unstable angina.

segment. The processing time of the MEM-based ECG recovery for one buffer is 2.55 *ms*. Therefore, the overall latency for our system is only 2.31 + 1.29 + 2.55 = 6.15 ms (< 10 ms), which is acceptable for real-time monitoring.

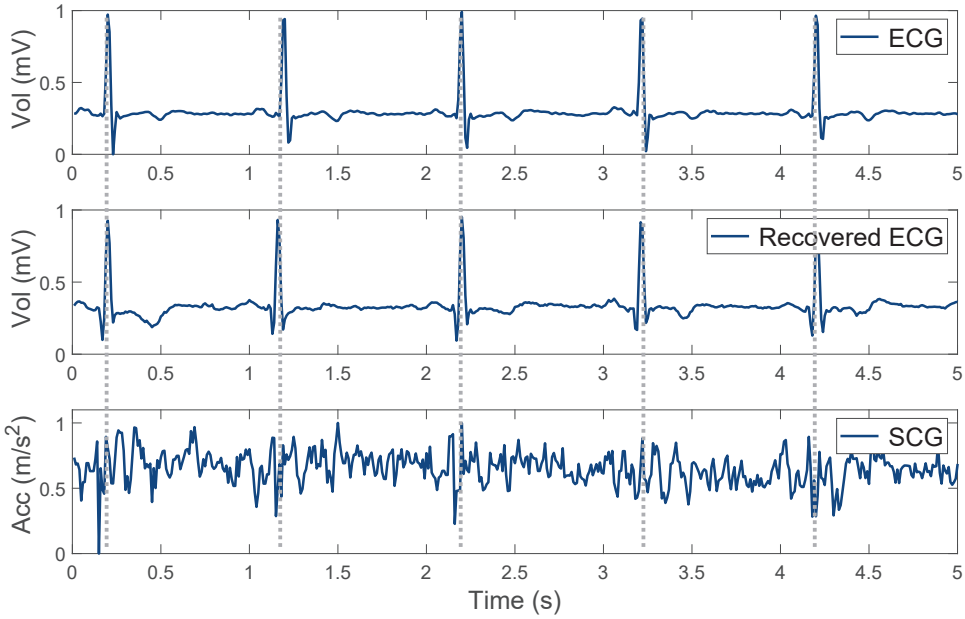


Fig. 25. Recovered ECG waveform from a patient with myocardial infarction.

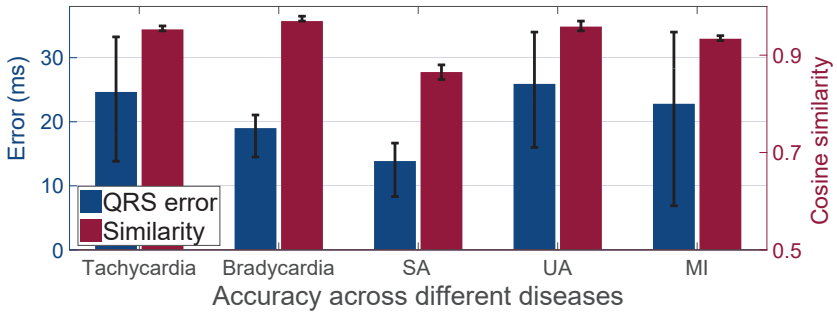


Fig. 26. Accuracy of ECG recovery for five cardiac diseases.

Our system exhibits moderate power consumption, with a measured average of 132.2 ± 7.3 mW when running on commercial mobile phones. We utilize PowerTutor [77] to explore the power usage of our system on a Samsung Galaxy S10. We measure the average power consumption over one hour, which is divided into six sessions, each consisting of six minutes. In each session, the user performs ten rounds of ECG recovery. We exclude the power consumption of the LCD display from our measurements, resulting in an average CPU power consumption of $132.2 \pm 7.3\ mW$.

8 RELATED WORK

We discuss the related work and group them into the following categories.

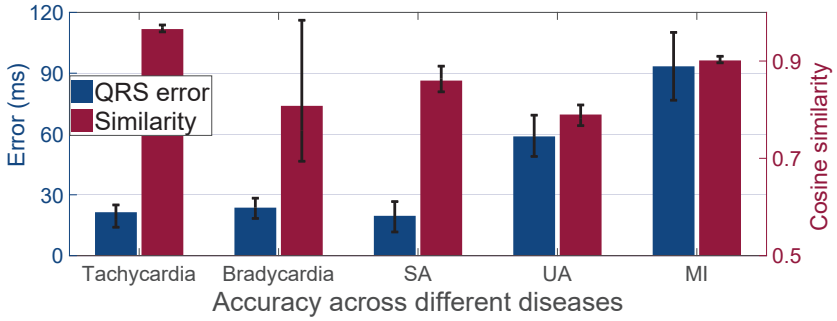


Fig. 27. Accuracy of ECG recovery for unseen diseases.

Device-based ECG monitoring: Recent years have seen a surge of interest in ECG monitoring as an important method for diagnosing cardiac disease. Currently, heartbeat monitoring is primarily achieved through wearable devices or sensors. Traditional commercial ECG devices [78] capture the electronic activity of the heart by attaching electrode pads to specific locations on the body. To facilitate ubiquitous ECG monitoring, several studies have focused on using flexible textile electrodes embedded in fabrics to monitor heart diseases in real-time [79, 80]. However, these schemes mainly rely on specialized equipment, which may not be suitable for the general public. With the increasing popularity of smartwatches, heart-related activity monitoring has become a standard feature. Companies such as Apple [81] and Samsung [82] have embedded electrodes in their devices to support cardiac monitoring. However, the current methods require a user to clip their finger to an ECG board to make contact with the electrodes, which may be inconvenient and uncomfortable. In addition, their accuracy is highly dependent on precise electrode location and proper contact. Recently, a study in [28] proposed using wristwatch vibrations to estimate ECG signals, but it relies on specialized equipment and experiences poor accuracy for new users who haven't been previously recorded. In comparison to all prior studies, we present a novel ECG sensing system that uses acceleration signals collected from a mobile device. Our system offers several advantages over previous methods, including almost zero cost, ease of use, location and user independence, and high accuracy.

Device-free ECG monitoring: Device-free ECG monitoring, which offers a comfortable user experience, has the potential for widespread adoption. Recent research has mainly focused on estimating ECG signals through wireless signals. Chen et al. [26] have developed a millimeter-wave radar system that can measure cardiac mechanical activity and reconstruct ECG signals without physical contact. Wang et al. [25] have employed a UWB device to achieve ECG-grained cardiac monitoring for the first time. Park et al. [27] have presented a low-cost, accurate, and non-intrusive geophone-based sensing system for extracting accurate ECG patterns using heartbeat vibrations that penetrate through a bed mattress. However, these fixed devices only work when users are located in a particular area, making it challenging to meet the requirement for daily, location-independent heartbeat monitoring. In contrast, our system is based on mobile phones, allowing for convenient heartbeat monitoring from any location.

9 LIMITATION AND DISCUSSION

We now discuss several limitations in our system. First, the current version of the system requires the user to remain still while collecting data due to the accelerometer's sensitivity to body movements. This restriction may limit the system's reliability in recovering the ECG waveform during activities such as running or walking. To address this limitation, we plan to explore contrastive learning techniques to extract heartbeat signals in future work. Second, at present, given the constrained computing resource of mobile phones, it's necessary to train deep learning models on computers. This approach, however, is not particularly favourable for real-time data and model

Table 2. Time consumption

	SWT-based Denoising	Soft-Thresholding Denoising	MEM based ECG Recovery	Total
Time	2.31 *ms*	1.29 *ms*	2.55 *ms*	6.15 *ms*

updates, thereby diminishing the system's robustness. In the future, we plan to proactively design lightweight deep learning models that can be trained directly on mobile devices so as to enable online model updates. Third, at present, our system has a relatively poor recovery capability for unseen cardiac diseases, especially those samples with drastic pattern changes in waveforms. In the future, we will consider other more advanced deep learning models to improve the recovery effect.

10 CONCLUSIONS

This paper introduces a novel low-cost ECG measurement system based on acceleration signals collected from smartphones. Our system offers several advantages over previous systems, including its low cost, ease of use, location independence, user independence, and high accuracy. To eliminate interference caused by respiration, hand movements, and system noise and increase the variability of the training data, we apply two-tiered denoising method. We also develop a multi-level deep learning recovery model to achieve fine-grained and user-independent ECG measurement. We have implemented the real-time ECG monitoring system using a single mobile phone and conducted extensive experiments demonstrating the strong capability of our system in ECG measurement, even for patients with various heart diseases.

ACKNOWLEDGMENTS

This research is supported by National Natural Science Foundation of China (Grant No. 62102006). This work is also in part supported by The Natural Science Foundation of the Jiangsu Higher Education Institutions of China (Grant No. 1020231697).

REFERENCES

[1] World Health Organization. Cardiovascular diseases (CVDs). https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds), 2021.
[2] Ancel Keys. Coronary heart disease—the global picture. *Atherosclerosis*, 22(2):149–192, 1975.
[3] Grant W Reed, Jeffrey E Rossi, and Christopher P Cannon. Acute myocardial infarction. *The Lancet*, 389(10065):197–210, 2017.
[4] George B Moody and Roger G Mark. The impact of the MIT-BIH arrhythmia database. *IEEE Engineering in Medicine and Biology Magazine*, 20(3):45–50, 2001.
[5] F Ter Woort, G Dubois, M Didier, and E Van Erck-Westergren. Validation of an equine fitness tracker: Heart rate and heart rate variability. *Comparative Exercise Physiology*, 17(2):189–198, 2021.
[6] Chelsea G Bender, Jason C Hoffstot, Brian T Combs, Sara Hooshangi, and Justin Cappos. Measuring the fitness of fitness trackers. In *2017 IEEE Sensors Applications Symposium (SAS)*, pages 1–6. IEEE, 2017.
[7] Qinge Wu, Kelli Sum, and Dan Nathan-Roberts. How fitness trackers facilitate health behavior change. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 2016.
[8] Tao Gu, Liang Wang, Hanhua Chen, Xianping Tao, and Jian Lu. Recognizing multiuser activities using wireless body sensor networks. *IEEE Transactions on Mobile Computing*, 10(11):1618–1631, 2011.
[9] 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.
[10] Ricard Delgado-Gonzalo, Jakub Parak, Adrian Tarniceriu, Philippe Renevey, Mattia Bertschi, and Ilkka Korhonen. Evaluation of accuracy and reliability of pulseon optical heart rate monitoring device. In *Proceedings of IEEE EMBC*, 2015.

[11] Liang Wang, Tao Gu, Hanhua Chen, Xianping Tao, and Jian Lu. Real-time activity recognition in wireless body sensor networks: From simple gestures to complex activities. *2010 IEEE 16th International Conference on Embedded and Real-Time Computing Systems and Applications*, pages 43–52, 2010.

[12] Dung Phan, Lee Yee Siong, Pubudu N Pathirana, and Aruna Seneviratne. Smartwatch: Performance evaluation for long-term heart rate monitoring. In *Proceedings of IEEE ISBB*, 2015.

[13] Nicole Morresi, Sara Casaccia, Matteo Sorcinelli, Marco Arnesano, and Gian Marco Revel. Analysing performances of heart rate variability measurement through a smartwatch. In *Proceedings of IEEE MeMeA*, 2020.

[14] Yu Zhang, Tao Gu, Chu Luo, Vassilis Kostakos, and Aruna Seneviratne. Findroidhr: Smartwatch gesture input with optical heartrate monitor. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2:1–42, 03 2018.

[15] Fatemeh Sarhaddi, Kianoosh Kazemi, Iman Azimi, Rui Cao, Hannakaisa Niela-Vilén, Anna Axelin, Pasi Liljeberg, and Amir M Rahmani. A comprehensive accuracy assessment of samsung smartwatch heart rate and heart rate variability. *PloS One*, 17(12):e0268361, 2022.

[16] Susanna Spinsante, Sara Porfiri, and Lorenzo Scalise. Accuracy of heart rate measurements by a smartwatch in low intensity activities. In *Proceedings of IEEE MeMeA*, 2019.

[17] Donald W Romhilt and E Harvey Estes Jr. A point-score system for the ECG diagnosis of left ventricular hypertrophy. *American Heart Journal*, 75(6):752–758, 1968.

[18] Selcan Kaplan Berkaya, Alper Kursat Uysal, Efnan Sora Gunal, Semih Ergin, Serkan Gunal, and M Bilginer Gulmezoglu. A survey on ECG analysis. *Biomedical Signal Processing and Control*, 43:216–235, 2018.

[19] Shlomo Stern, D Tzivoni, and Z Stern. Diagnostic accuracy of ambulatory ECG monitoring in ischemic heart disease. *Circulation*, 52(6):1045–1049, 1975.

[20] Nicolas Chin-Yee, Gianni D'Egidio, Kednapa Thavorn, Daren Heyland, and Kwadwo Kyeremanteng. Cost analysis of the very elderly admitted to intensive care units. *Critical Care*, 21(1):1–7, 2017.

[21] BPL 6208 View, 3 channel ECG Machine. https://www.hospitalsstore.com/bpl-6208-ecg-machine-price/.

[22] Marc Strik, Sylvain Ploux, Daniel Weigel, Joske van der Zande, Anouk Velraeds, Hugo-Pierre Racine, F Daniel Ramirez, Michel Haïssaguerre, and Pierre Bordachar. The use of smartwatch electrocardiogram beyond arrhythmia detection. *Trends in Cardiovascular Medicine*, 2023.

[23] Amirali Behzadi, Alireza Sepehri Shamloo, Konstantinos Mouratis, Gerhard Hindricks, Arash Arya, and Andreas Bollmann. Feasibility and reliability of smartwatch to obtain 3-lead electrocardiogram recordings. *Sensors*, 20(18):5074, 2020.

[24] Mathieu Nasarre, Marc Strik, Francisco Daniel Ramirez, Samuel Buliard, Hugo Marchand, Saer Abu-Alrub, Sylvain Ploux, Michel Haïssaguerre, and Pierre Bordachar. Using a smartwatch electrocardiogram to detect abnormalities associated with sudden cardiac arrest in young adults. *EP Europace*, 24(3):406–412, 2022.

[25] Zhi Wang, Beihong Jin, Siheng Li, Fusang Zhang, and Wenbo Zhang. ECG-grained cardiac monitoring using uwb signals. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 6(4):1–25, 2023.

[26] Jinbo Chen, Dongheng Zhang, Zhi Wu, Fang Zhou, Qibin Sun, and Yan Chen. Contactless electrocardiogram monitoring with millimeter wave radar. *IEEE Transactions on Mobile Computing*, 2022.

[27] Jaeyeon Park, Hyeon Cho, Rajesh Krishna Balan, and JeongGil Ko. Heartquake: Accurate low-cost non-invasive ECG monitoring using bed-mounted geophones. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 4(3):1–28, 2020.

[28] Yetong Cao, Fan Li, Huijie Chen, Xiaochen liu, Li Zhang, and Yu Wang. Guard your heart silently: Continuous electrocardiogram waveform monitoring with wrist-worn motion sensor. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 6(3):1–29, 2022.

[29] Amer Abdulmahdi Chlaihawi, Binu Baby Narakathu, Sepehr Emamian, Bradley J Bazuin, and Massood Z Atashbar. Development of printed and flexible dry ECG electrodes. *Sensing and Bio-sensing Research*, 20:9–15, 2018.

[30] Leonard S Lilly. *Pathophysiology of heart disease: a collaborative project of medical students and faculty*. Lippincott Williams & Wilkins, 2012.

[31] Jingting Yao, Srini Tridandapani, Carson A Wick, and Pamela T Bhatti. Seismocardiography-based cardiac computed tomography gating using patient-specific template identification and detection. *IEEE Journal of Translational Engineering in Health and Medicine*, 5:1–14, 2017.

[32] Prasan Kumar Sahoo, Hiren Kumar Thakkar, Wen-Yen Lin, Po-Cheng Chang, and Ming-Yih Lee. On the design of an efficient cardiac health monitoring system through combined analysis of ECG and SCG signals. *Sensors*, 18(2):379, 2018.

[33] Seyed Ali Khonsary. Guyton and hall: textbook of medical physiology. *Surgical Neurology International*, 8, 2017.

[34] Arnold M Katz. *Physiology of the Heart*. Lippincott Williams & Wilkins, 2010.

[35] Mohammad Shamim Imtiaz, Rajeena Shrestha, Talwinder Dhillon, Kazi Ata Yousuf, Bilal Saeed, Anh Dinh, and Khan Wahid. Correlation between seismocardiogram and systolic blood pressure. In *Proceedings of IEEE CCECE*, 2013.

[36] M Di Rienzo, E Vaini, P Castiglioni, G Merati, P Meriggi, G Parati, A Faini, and F Rizzo. Wearable seismocardiography: Towards a beat-by-beat assessment of cardiac mechanics in ambulant subjects. *Autonomic Neuroscience*, 178(1-2):50–59, 2013.

[37] Richard S Crow, Peter Hannan, David Jacobs, Lowell Hedquist, and David M Salerno. Relationship between seismocardiogram and echocardiogram for events in the cardiac cycle. *American Journal of Noninvasive Cardiology*, 8:39–46, 1994.
[38] Gustavo Lenis, Nicolas Pilia, Axel Loewe, Walther HW Schulze, and Olaf Dössel. Comparison of baseline wander removal techniques considering the preservation of st changes in the ischemic ECG: a simulation study. *Computational and Mathematical Methods in Medicine*, 2017.
[39] Shadnaz Asgari and Alireza Mehrnia. A novel low-complexity digital filter design for wearable ECG devices. *PloS One*, 12(4):e0175139, 2017.
[40] Zhongyi Jin, Anming Dong, Minglei Shu, and Yinglong Wang. Sparse ECG denoising with generalized minimax concave penalty. *Sensors*, 19(7):1718, 2019.
[41] Ahilan Appathurai, J Jerusalin Carol, C Raja, SN Kumar, Ashy V Daniel, A Jasmine Gnana Malar, A Lenin Fred, and Sujatha Krishnamoorthy. A study on ECG signal characterization and practical implementation of some ECG characterization techniques. *Measurement*, 147:106384, 2019.
[42] Manas Rakshit and Susmita Das. An efficient ECG denoising methodology using empirical mode decomposition and adaptive switching mean filter. *Biomedical Signal Processing and Control*, 40:140–148, 2018.
[43] Shailesh Kumar, Damodar Panigrahy, and PK Sahu. Denoising of electrocardiogram (ECG) signal by using empirical mode decomposition (emd) with non-local mean (nlm) technique. *Biocybernetics and Biomedical Engineering*, 38(2):297–312, 2018.
[44] Mohammad Reza Mohebbian, Mohammad Wajih Alam, Khan A Wahid, and Anh Dinh. Single channel high noise level ECG deconvolution using optimized blind adaptive filtering and fixed-point convolution kernel compensation. *Biomedical Signal Processing and Control*, 57:101673, 2020.
[45] Ashish Kumar, Harshit Tomar, Virender Kumar Mehla, Rama Komaragiri, and Manjeet Kumar. Stationary wavelet transform based ECG signal denoising method. *ISA Transactions*, 114:251–262, 2021.
[46] Arthur C Guyton, John Edward Hall, et al. *Textbook of medical physiology*, volume 548. Saunders Philadelphia, 1986.
[47] Yetong Cao, Qian Zhang, Fan Li, Song Yang, and Yu Wang. PPGPass: Nonintrusive and secure mobile two-factor authentication via wearables. In *Proceedings of IEEE Infocom*, 2020.
[48] 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 the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 7(1):1–24, 2023.
[49] David L Donoho. De-noising by soft-thresholding. *IEEE Transactions on Information Theory*, 41(3):613–627, 1995.
[50] Jeena Joy, Salice Peter, and Neetha John. Denoising using soft thresholding. *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, 2(3):1027–1032, 2013.
[51] PM Agante and JP Marques De Sá. ECG noise filtering using wavelets with soft-thresholding methods. In *Computers in Cardiology 1999. Vol. 26 (Cat. No. 99CH37004)*, pages 535–538. IEEE, 1999.
[52] Vatsal Gajera, Rishabh Gupta, Prasanta K Jana, et al. An effective multi-objective task scheduling algorithm using min-max normalization in cloud computing. In *Proceedings of IEEE iCATccT*, 2016.
[53] Chung-Cheng Chiu, Tara N Sainath, Yonghui Wu, Rohit Prabhavalkar, Patrick Nguyen, Zhifeng Chen, Anjuli Kannan, Ron J Weiss, Kanishka Rao, Ekaterina Gonina, et al. State-of-the-art speech recognition with sequence-to-sequence models. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 4774–4778. IEEE, 2018.
[54] Dalin Zhang, Lina Yao, Kaixuan Chen, Zheng Yang, Xin Gao, and Yunhao Liu. Preventing sensitive information leakage from mobile sensor signals via integrativetransformation. *IEEE Transactions on Mobile Computing*, 2021.
[55] Sixiao Zheng, Jiachen Lu, Hengshuang Zhao, Xiatian Zhu, Zekun Luo, Yabiao Wang, Yanwei Fu, Jianfeng Feng, Tao Xiang, Philip HS Torr, et al. Rethinking semantic segmentation from a sequence-to-sequence perspective with transformers. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 6881–6890, 2021.
[56] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International Conference on Machine Learning*, pages 448–456. pmlr, 2015.
[57] Andrew L Maas, Awni Y Hannun, Andrew Y Ng, et al. Rectifier nonlinearities improve neural network acoustic models. In *International Conference on Machine Learning*, pages 3–10. Atlanta, Georgia, USA, 2013.
[58] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L Yuille. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(4):834–848, 2017.
[59] Peter J Huber. Robust estimation of a location parameter. *Breakthroughs in Statistics: Methodology and Distribution*, pages 492–518, 1992.
[60] Ross Girshick. Fast r-cnn. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pages 1440–1448, 2015.
[61] Jia-Ren Chang and Yong-Sheng Chen. Pyramid stereo matching network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5410–5418, 2018.

[62] Martin Brossard, Silvere Bonnabel, and Axel Barrau. Denoising imu gyroscopes with deep learning for open-loop attitude estimation. *IEEE Robotics and Automation Letters*, 5(3):4796–4803, 2020.
[63] Jiqiang Wang, Baohu Wu, Peng Wei, Shengtong Sun, and Peiyi Wu. Fatigue-free artificial ionic skin toughened by self-healable elastic nanomesh. *Nature Communications*, 13(1):4411, 2022.
[64] Fusang Zhang, Zhi Wang, Beihong Jin, Jie Xiong, and Daqing Zhang. Your smart speaker can "hear" your heartbeat! *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 4(4):1–24, 2020.
[65] Kun Qian, Chenshu Wu, Fu Xiao, Yue Zheng, Yi Zhang, Zheng Yang, and Yunhao Liu. Acousticcardiogram: Monitoring heartbeats using acoustic signals on smart devices. In *2018 IEEE Conference on Computer Communications*, pages 1574–1582. IEEE, 2018.
[66] Tianyue Zheng, Zhe Chen, Chao Cai, Jun Luo, and Xu Zhang. V2iFi: In-vehicle vital sign monitoring via compact rf sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 4(2):1–27, 2020.
[67] Portable ECG Monitors. Heal Force PC-80B. http://www.healforce.com/en/html/products/portableecgmonitors/healthcare-equipment-portable-ECG-monitors-PC-80B.html.
[68] Michael B Simson. Use of signals in the terminal qrs complex to identify patients with ventricular tachycardia after myocardial infarction. *Circulation*, 64(2):235–242, 1981.
[69] Amir Kashani and S Serge Barold. Significance of qrs complex duration in patients with heart failure. *Journal of the American College of Cardiology*, 46(12):2183–2192, 2005.
[70] Borys Surawicz and Timothy Knilans. *Chou's electrocardiography in clinical practice: adult and pediatric*. Elsevier Health Sciences, 2008.
[71] Fred M Kusumoto, Mark H Schoenfeld, Coletta Barrett, James R Edgerton, Kenneth A Ellenbogen, Michael R Gold, Nora F Goldschlager, Robert M Hamilton, José A Joglar, Robert J Kim, et al. 2018 acc/aha/hrs guideline on the evaluation and management of patients with bradycardia and cardiac conduction delay: a report of the american college of cardiology/american heart association task force on clinical practice guidelines and the heart rhythm society. *Journal of the American College of Cardiology*, 74(7):e51–e156, 2019.
[72] Jay W Mason, Douglas J Ramseth, Dennis O Chanter, Thomas E Moon, Daniel B Goodman, and Boaz Mendzelevski. Electrocardiographic reference ranges derived from 79,743 ambulatory subjects. *Journal of Electrocardiology*, 40(3):228–234, 2007.
[73] Michael O Sweeney, Rutger J van Bommel, Martin J Schalij, C Jan Willem Borleffs, Anne S Hellkamp, and Jeroen J Bax. Analysis of ventricular activation using surface electrocardiography to predict left ventricular reverse volumetric remodeling during cardiac resynchronization therapy. *Circulation*, 121(5):626–634, 2010.
[74] Biagio Sassone, Simona Gambetti, Matteo Bertini, Matteo Beltrami, Giosuè Mascioli, Sabrina Bressan, Giuseppe Fucà, Federico Pacchioni, Mario Pedaci, Federica Michelotti, et al. Relation of qrs duration to response to cardiac resynchronization therapy. *The American Journal of Cardiology*, 115(2):214–219, 2015.
[75] Paul Kligfield, Leonard S Gettes, James J Bailey, Rory Childers, Barbara J Deal, E William Hancock, Gerard Van Herpen, Jan A Kors, Peter Macfarlane, David M Mirvis, et al. Recommendations for the standardization and interpretation of the electrocardiogram: part i: the electrocardiogram and its technology: a scientific statement from the american heart association electrocardiography and arrhythmias committee, council on clinical cardiology; the american college of cardiology foundation; and the heart rhythm society endorsed by the international society for computerized electrocardiology. *Circulation*, 115(10):1306–1324, 2007.
[76] Ellen C Keeley and L David Hillis. Primary pci for myocardial infarction with st-segment elevation. *New England Journal of Medicine*, 356(1):47–54, 2007.
[77] Z Yang. Powertutor-a power monitor for android-based mobile platforms. *EECS, University of Michigan, retrieved September*, 2:19, 2012.
[78] Holter monitor. https://www.mayoclinic.org/tests-procedures/holter-monitor/about/pac-20385039., 2022.
[79] Nazmul Karim, Shaila Afroj, Andromachi Malandraki, Sean Butterworth, Christopher Beach, Muriel Rigout, Kostya S Novoselov, Alexander J Casson, and Stephen G Yeates. All inkjet-printed graphene-based conductive patterns for wearable e-textile applications. *Journal of Materials Chemistry C*, 5(44):11640–11648, 2017.
[80] Merja M Puurtinen, Satu M Komulainen, Pasi K Kauppinen, Jaakko AV Malmivuo, and Jari AK Hyttinen. Measurement of noise and impedance of dry and wet textile electrodes, and textile electrodes with hydrogel. In *2006 International Conference of the IEEE Engineering in Medicine & Biology Society*, pages 6012–6015. IEEE, 2006.
[81] Record an ECG on apple watch. https://support.apple.com/en-us/HT208955., 2022.
[82] ECG at your fingertips. https://www.samsung.com/us/apps/samsung-health-monitor/., 2022.