



# **AirECG: Contactless Electrocardiogram for Cardiac Disease Monitoring via mmWave Sensing and Cross-domain Diffusion Model**

**LANGCHENG ZHAO, RUI LYU, HANG LEI, and QI LIN**, State key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, China

**ANFU ZHOU\*** and **HUADONG MA\***, State key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, China

**JINGJIA WANG, XIANGBIN MENG, CHUNLI SHAO, and YIDA TANG**, Department of Cardiology, Peking University Third Hospital, China

**GUOXUAN CHI and ZHENG YANG**, School of Software and BNRist, Tsinghua University, China

The electrocardiogram (ECG) has always served as a crucial biomedical examination for cardiac diseases monitoring and diagnosing. Typical ECG measurement requires attaching electrodes to the body, which is inconvenient for long-term monitoring. Recent wireless sensing maps wireless signals reflected from human chest into electrical activities of heart so as to reconstruct ECG contactlessly. While making great progress, we find existing works are effective only for healthy populations with normal ECG, but fall short when confronted with the most desired usage scenario: reconstructing ECG accurately for people with cardiac diseases such as atrial fibrillation, premature ventricular beat. To bridge the gap, we propose *AirECG*, which moves forward to reconstruct ECG for both healthy people and even cardiac patients with morbid ECG, *i.e.*, irregular rhythm and anomalous ECG waveform, via contactless millimeter-wave sensing. To realize *AirECG*, we first custom-design a cross-domain diffusion model that can perform multiple iteration denoising inference, in contrast with the single-step generative models widely used in previous works. In this way, *AirECG* is able to identify and eliminate the distortion due to the unstable and irregular cardiac activities, so as to synthesize ECG even during abnormal beats. Furthermore, we enhance the determinacy of *AirECG*, *i.e.*, to generate high-fidelity ECG, by designing a calibration guidance mechanism to combat the inherent randomness issue of the probabilistic diffusion model. Empirical evaluation demonstrates *AirECG*'s ability of ECG synthesis with Pearson correlation coefficient (PCC) of 0.955 for normal beats. Especially for abnormal beats, the PCC still exhibits a strong correlation of 0.860, with 15.0%~21.1% improvement compared with state-of-the-art approaches.

CCS Concepts: • **Human-centered computing** → **Ubiquitous and mobile computing systems and tools**; *Ubiquitous and mobile computing design and evaluation methods*; *Empirical studies in ubiquitous and mobile computing*.

Additional Key Words and Phrases: Contactless ECG monitoring, millimeter wave sensing, diffusion model

\*Corresponding author: Anfu Zhou and Huadong Ma.

---

Authors' Contact Information: **Langcheng Zhao**, zhaolangcheng@bupt.edu.cn; **Rui Lyu**, lvrui1999@bupt.edu.cn; **Hang Lei**, 1930422631@qq.com; **Qi Lin**, linky@bupt.edu.cn, State key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing, China; **Anfu Zhou**, zhouanfu@bupt.edu.cn; **Huadong Ma**, mhd@bupt.edu.cn, State key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing, China; **Jingjia Wang**, wjingjia@126.com; **Xiangbin Meng**, 15896850171@163.com; **Chunli Shao**, chunlishao@126.com; **Yida Tang**, tangyida@bjmu.edu.cn, Department of Cardiology, Peking University Third Hospital, Beijing, China; **Guoxuan Chi**, chiguoxuan@gmail.com; **Zheng Yang**, hmilyyz@gmail.com, School of Software and BNRist, Tsinghua University, Beijing, China.

---

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, or post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM 2474-9567/2024/9-ART144

<https://doi.org/10.1145/3678550>

**ACM Reference Format:**

Langcheng Zhao, Rui Lyu, Hang Lei, Qi Lin, Anfu Zhou, Huadong Ma, Jingjia Wang, Xiangbin Meng, Chunli Shao, Yida Tang, Guoxuan Chi, and Zheng Yang. 2024. *AirECG*: Contactless Electrocardiogram for Cardiac Disease Monitoring via mmWave Sensing and Cross-domain Diffusion Model. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 8, 3, Article 144 (September 2024), 27 pages. <https://doi.org/10.1145/3678550>

**1 Introduction**

Cardiovascular diseases (CVDs) are the most critical health issue in the world, they take an estimated 17.9 million lives per year and even show a growing trend [3, 35]. The Electrocardiogram (ECG) is a common and painless test used to detect types of CVDs and monitor the heart's health, which greatly helps disease management [5, 33]. Specifically, daily ECG monitoring and early diagnosis play important roles in controlling disease progression and preventing adverse events [10]. Taking the most common arrhythmia as an example of CVDs, with the help of daily ECG tests and effective clinical treatment like anticoagulation, serious complications like stroke and even mortality can be prevented [26].

Due to the urgency and great demand for ECG monitoring, relevant technology development is high-profile, particularly portable ECG devices that can achieve disease monitoring and management. As the general devices for ECG examination, standard electrocardiograph and desktop ECG monitors are widely used, while they are bulky and require electrodes attached to the skin. Wearable Holter [6] and ECG patches [4, 45] solve the problem of device size, however, they still face negative issues from electrodes. Those electrodes may lead to unsatisfactory experience and skin allergy, which hinders the daily ECG monitoring usage. Besides, the electrode attachment limits ECG monitoring for specific patients, such as neonates and patients with skin damage. Newly released smartwatches are equipped with a single lead ECG sensor, which requires the user to actively touch the electrode button during ECG measurement, nevertheless such a design is not suitable for continuous ECG monitoring [24].

To achieve daily cardiac monitoring with passive and continuous advantages, wireless sensing solutions have been explored to monitor the thorax wall vibration using WiFi [40, 41], RFID [30], and millimeter wave (mmWave) radar [20, 39]. These works monitor the mechanical activities of chest caused by heartbeats, so as to measure simple heart rate and its variability. Furthermore, latest works attempt to transfer the radar signal into heart electrical activities as ECG [11, 42, 43], preliminarily achieving contactless ECG monitoring. Existing methods are based on single-step generative neural networks, and perform reasonable ECG for users with normal heartbeats. While for patients with morbid heartbeats, existing generative models may lose efficacy and lead to ECG distortion. For instance, abnormal heartbeats of atrial fibrillation (AF) lead to invalid ECG generation without cardiac features, resulting in disability to detect disease. To sum up, the contactless ECG for disease monitoring is still challenging due to the abnormal heartbeats caused by diseases.

To bridge the gap, we propose *AirECG*, a contactless ECG measurement system that enables passive and continuous disease monitoring for both healthy and cardiac patients in daily life. Fig. 1 illustrates the application scenarios of *AirECG*, where we set up a low-cost mmWave radar [8] to sense the subject's thorax wall vibration through mmWave propagated in the air. Then we custom-designed a cross-domain diffusion model to transfer the heart vibration in mechanical domain into ECG electrical domain, and our model provides faithful domain transformation through the multiple denoising iterations and calibration guidance. In this way, *AirECG* can provide an accurate ECG even during diseases like atrial fibrillation, premature beat, etc., so as to offer early warnings and better treatment management. To realize *AirECG*, we identify and address two main challenges as follows.

**Modeling scarce and irregular morbid cardiac motions by cross-domain diffusion generative model.**

Different from previous works focused on normal and healthy heartbeats, *AirECG* concerns investigating the contactless ECG monitoring during abnormal beats. For healthy subjects, the inter-beat intervals of consecutive beats are similar, and the cardiac activities inside each beat are the same (regular contraction and relaxation of atria

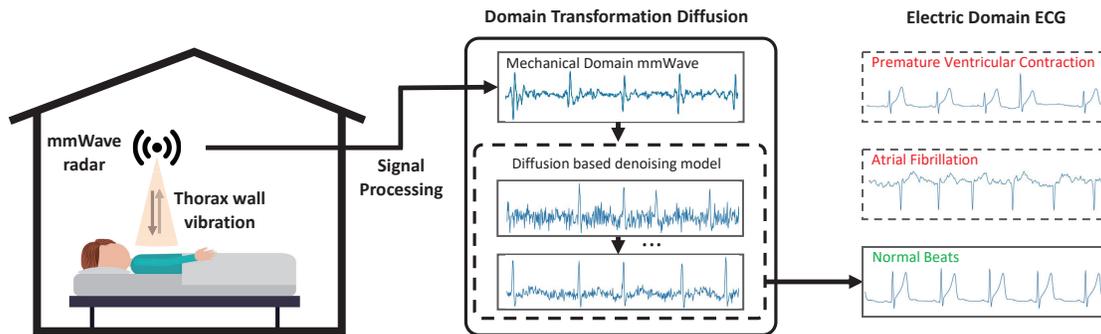


Fig. 1. The usage scenarios of *AirECG*. *AirECG* can provide contactless ECG for disease monitoring, which is based on domain transformation from mmWave mechanical domain to electrical domain ECG.

and ventricles). However, for patients with arrhythmia such as AF, the heartbeats become chaotic and irregular [2], which makes the existing contactless ECG performance decline. The main reason is that sophisticated ECG during arrhythmia can not be directly transformed from mmWave signal using single-step models like TCN (Temporal Convolutional Networks) or GAN (Generative Adversarial Network) in previous work. Therefore, we propose a cross-domain diffusion model, which can perform multiple iteration inference so as to synthesize and revise electrical domain ECG based mechanical domain mmWave data. Despite the fact that abnormal heartbeats are hard to synthesize directly, cross-domain diffusion can revise the incomplete ECG from the last iteration. In particular, cross-domain diffusion is initialized from Gaussian noise and performs step-by-step denoising inference based on mmWave. With the help of multiple-step revision deeply combined with mmWave, *AirECG* can synthesize ECG even during irregular morbid heartbeats.

**Enhancing ECG synthesis deterministic against generative model's randomness nature.** Both our cross-domain diffusion model and previous ECG synthesis models are types of generative models. The randomness of generative models arises from the principles of probabilistic design, such as the Gaussian initialization and denoising procedure of diffusion models. For content generation in the field of image or text generation, the ability to freely exert is welcomed to improve the diversity of generated content. However, when it comes to ECG synthesis, the fidelity of model output is required. To this end, we design a neural network plug-in unit called ECG calibration guidance, which controls the random denoise step of diffusion model based on cross-attention guidance. Our design utilizes previous ECG waveforms from calibration device, thus it can provide personal calibration information for the denoising process, avoiding output distortion when dealing with new subjects apart from training data. In this way, the accuracy of contactless ECG monitoring can be further improved, especially for abnormal heartbeats and unseen subjects.

We prototype *AirECG* on Texas Instruments IWR1443BOOST COTS mmWave radar [7, 8], while using the ECG patch [4] as the calibration device and ground truth for training and evaluating *AirECG*. For datasets collection, we recruit practical participants with medical professionals' cooperation, then collect mmWave data with typical types of cardiac diseases, e.g., atrial fibrillation (AF), premature ventricular beat (PVC). We also perform 5-fold user-dependent and user-independent validation on 109,598 heartbeats to in-depthly verify the performance of *AirECG*. Overall, *AirECG* achieves compelling performance, showing a Pearson correlation coefficient (PCC) of 0.955 for normal beats. Especially for abnormal beats, the PCC still exhibits a strong correlation of 0.860, with 15.0%~21.1% improvement compared with state-of-the-art approaches. We further validate the arrhythmia detection performance as a downstream application, which shows 98.82% accuracy and 0.97 F1-score to discriminate normal or abnormal heartbeat.

**Contributions:** Our contribution can be summarized as follows,

- We propose a novel cross-domain diffusion model, which transfers mmWave signal modulated by cardiac motion in mechanical domain, into ECG waveform in electrical domain.
- We design an ECG calibration guidance plug-in based on cross-attention module, so as to avoid free generation of diffusion model, and further improve ECG synthesis fidelity.
- We prototype and validate *AirECG* system with multiple types of arrhythmia diseases, through the collaboration of medical professionals. To our knowledge, this is the first-of-its-kind field test of contactless ECG evaluation with arrhythmia patients.

To contribute to the IMWUT community, we will update our source code and sample data at <https://github.com/LangchengZhao/AirECG>.

## 2 Understanding Contactless ECG Monitoring

In this section, we first introduce the basics of heart activities and principles of contactless ECG with mmWave radar. Then we examine the practical challenges of contactless ECG and find that existing ECG generation approaches lose efficacy and lead to ECG distortion in cases of arrhythmia patients exhibiting irregular morbid beats.

### 2.1 Wireless Cardiac Sensing

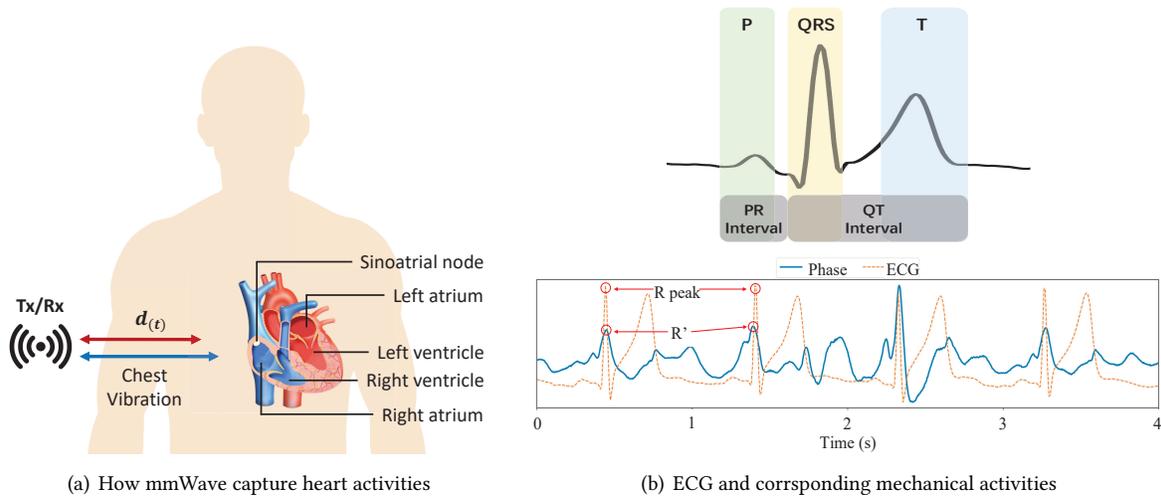


Fig. 2. Chest vibration changes the phase of mmWave signal. (a) is the sketch of mmWave cardiac sensing, (b) is an example of mmWave phase, filtered cardiac phase, and corresponding ECG ground truth. The mmWave filtered phase shows similar periodicity with ground truth, proving the feasibility of mmWave cardiac sensing.

A human heart is approximately 120mm in length, 80mm in width, and 60mm in thickness [38]. The heart can be divided into four chambers as Fig. 2(a) shows: the left and right atria (upper half), and the left and right ventricles (lower half). The atria are responsible for receiving the blood flowing back to the heart and pumping it into ventricles, while the ventricles can store the blood and pump it throughout the body. The mechanical activity of the heart is stimulated by the electrical activity of the heart, while the gold standard in medical practice is the electrical activity, which is described in the form of Electrocardiogram (ECG). The upper half of Fig. 2(b) shows

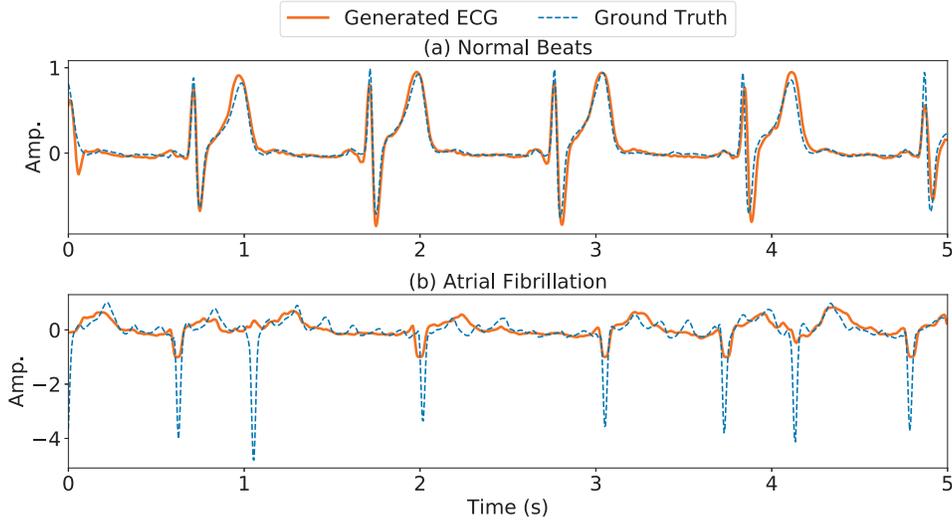


Fig. 3. Examples of generated contactless ECG from mmWave via existing method. (a) shows the ECG generation with high accuracy during normal heartbeats. (b) is the ECG generation when arrhythmia (atrial fibrillation) occurs, showing limited signal fidelity.

an ECG signal example of single cardiac cycle, where P wave represents atria depolarization, QRS complex for ventricular depolarization and T wave for ventricular re-polarization. It should be noted that atria re-polarization is covered by ventricular depolarization in QRS complex.

Fig. 2(a) illustrates the sketch of wireless cardiac sensing. In a normal cardiac cycle, the atria and ventricles contract sequentially and regularly, inducing vibration of 0.2-0.5mm in different parts of the thorax wall [16, 25]. Such vibration can be captured by mmWave radar, which emits mmWave to the thorax wall and receives the reflected signal that contains cardiac vibration information. We define the thorax wall vibration as  $d(t)$ , which can be extracted from the phase  $\phi_t$  of reflected mmWave signal as Equation. 1, where  $\lambda$  is mmWave wavelength.

$$d(t) = \frac{\lambda * \phi_t}{2\pi} \quad (1)$$

As mentioned above, mmWave radar records the mechanical activities of hearts as  $d(t)$ . The mechanical activity of the heart is generated under the stimulation of electrical activity, *i.e.*, the mechanical contraction of heart is initiated by electrical depolarization [32]. To verify the relationship between electrical and mechanical activities, we plot the synchronous mmWave phase  $\phi_t$  and ECG of a 4s segment in Fig. 2(b). The orange dashed line is the ECG captured by ECG patch [4], and the blue line plots the phase of mmWave. Since vibration caused by respiration is much greater than that of heartbeats, the phase of mmWave is processed with band pass filter to eliminate respiration and thereby shows the heart activity. We can find the peaks of heartbeats noted as  $R'$ , that corresponds to the ventricle depolarization R peak of ECG. Either  $R'$  or R shows peaks in 4s, *i.e.*, heart rate at 60beats/min. To sum up, the consistent cardiac cycle between mmWave and ECG can prove the correspondence between electrical and mechanical activities of the heart, which brings the feasibility of contactless cardiac monitoring.

## 2.2 Limitations of Contactless ECG Monitoring during Arrhythmia

To achieve contactless ECG monitoring, existing methods CardiacWave [43], CTL-ECG [11] or RF-ECG [42] transfer wireless signal into electric domain ECG signal. These methods utilize LSTM, TCN or cGAN (Conditional

GAN) models respectively for ECG generation, in particular, taking wireless signals as input and generating the ECG-like signals through one-step or autoregressive (one sample point each time) inference. These methods for ECG generation perform well during regular heartbeats with healthy people. We implement the latest contactless RF-ECG[42] using mmWave radar, where mmWave works as conditional information for ECG generation by cGAN. Fig. 3(a) shows a 5-second period of contactless ECG during normal heartbeats in orange line, while the blue dashed line plots the ground truth from ECG patch. The consistent relation between two ECG waves shows qualified performance with normal heart beats.

During normal heartbeats, the electrical impulse is produced by sinoatrial node of the heart in Fig. 2(a), and the impulse travels to atria and ventricles sequentially for depolarization, resulting in regular P-wave and QRS complex respectively. However, during atrial fibrillation (AF), *i.e.*, when completely irregular heartbeats occur, the ECG changes a lot. Specifically, as shown in the ground truth of Fig. 3(b), abnormal electrical impulses are wrongly and irregularly produced by atria, thus rhythm and amplitude of QRS-complex become irregular. Besides, normal P-waves disappear and are replaced by tremor waves (f-waves) of varying sizes and shapes. In such conditions, we found existing solutions show a performance degradation. As Fig. 3(b) shows, the cGAN model can capture the vast majority of heartbeats from mmWave, and attempts to synthesize them into ECG waves. While we find that the synthesized ECG is not well fitted with the ground truth ECG, *i.e.*, the generated ECG in orange line exhibits inconsistent cardiac features compared with blue dashed ground truth. In particular, (i) the morphology of each heartbeat exhibits low amplitude in the QRS-complex, and the details of f-waves disappear. (ii) Some heartbeats fail to be synthesized, such as the second and penultimate beats. Thus the rhythm and morphology of ECG are inconsistent between generated ECG and ground truth, leading to inadequate signal fidelity for disease monitoring.

The reason for such phenomenon is the sophisticated and irregular cardiac activities during AF, which make it difficult to realize mechanical-electrical domain transformation through prior models. In particular, for prior LSTM, TCN or cGAN models, sample points of ECG are generated in a single step without revision. However, when facing sophisticated synthesis like AF, these models lead to performance decline due to insufficient generative ability. The limitations of single step generative models can be summarized as, (i) limited generalization capability caused by the direct training objective from source domain to target domain, which has negative performance impacts on varying data distribution of test datasets. Especially for AF diseases, the test data varies a lot compared with training datasets, such as the rhythm, f-wave frequency. (ii) When the model inference is in single step, it means that the output cannot be augmented or revised any more.

For sophisticated content synthesis, the denoising diffusion probabilistic model (DDPM) [23] has shown its potential in image [17, 23] and audio generation [12]. Multiple iteration denoising has been proven [23] to be able to generate more sophisticated and irregular distributions in the image field. In recent studies, the feasibility of sequential signal generation and domain adaption has been validated, *i.e.*, RF-Diffusion [14] is the first generative diffusion model for RF signals based on a novel time-frequency diffusion theory, XFall [15] proposes a novel domain-generalized model design for Wi-Fi-based fall detection. In this paper, we aim to custom-design a diffusion-based model to realize sophisticated mechanical-electrical domain transformation, so as to achieve contactless ECG for cardiac disease monitoring. The advantages of multi step diffusion against single step models are in two aspects, (i) The generalization capability improves since the training objective is optimized into multiple iteration denoising rather than direct synthesis, which will be further discussed in Sec. 7.2 (ii) With the help of more inference steps as augmentation, the synthesis ability can increase gradually as discussed later in Sec. 6.5.

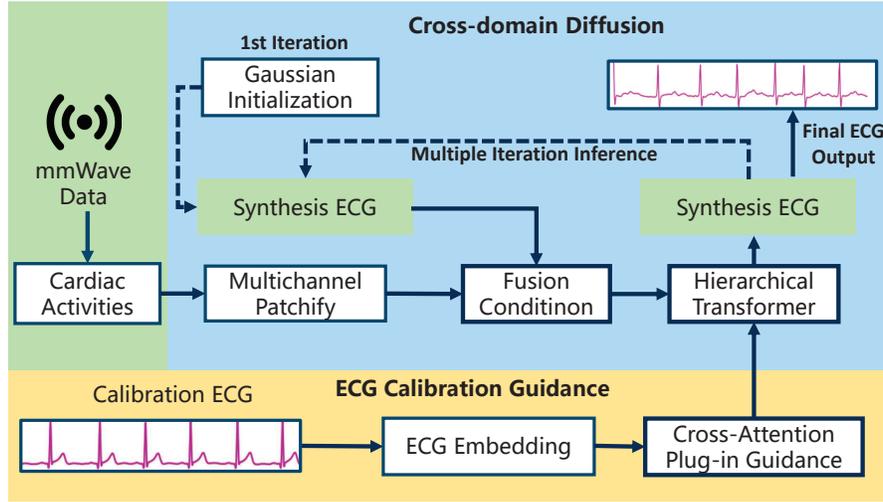


Fig. 4. The overview diagram of *AirECG*, which captures the cardiac activities through mmWave radar and then transfers them into ECG through cross-domain diffusion model. Additionally, an ECG calibration guidance module is proposed to eliminate randomness of generative process, thereby supporting faithful and high accuracy contactless ECG.

### 3 System Overview

In this section, we give *AirECG*'s system overview. *AirECG* focuses on contactless ECG for disease monitoring using mmWave radar signals. Fig. 4 illustrates the overview of *AirECG*, which consists of *cross-domain diffusion* and *ECG calibration guidance*.

*AirECG* captures the mechanical activities of hearts from mmWave radar, and then transfers them into ECG signals in electrical domain through cross-domain model. Specifically, regarding the workflow of *AirECG*, the mmWave signal processing is firstly realized to capture thorax wall vibration and extract cardiac mechanical activities. Then for the core design, we propose a cross-domain diffusion model so as to transfer mechanical domain activities into electrical domain ECG, which is realized through multiple iteration denoising inference. However, the denoising diffusion model is negatively impacted by the randomness of generative process, resulting in insufficient ECG signal fidelity. To solve this issue, we propose an ECG calibration guidance module, which guides the ECG synthesis by the historical ECG from reference device, supporting faithful and high accuracy contactless ECG for disease monitoring.

It is worth pointing out that, compared to previous works, *AirECG* focuses on contactless ECG for real disease data, and builds a customized cross-domain diffusion model to perform multiple iteration ECG synthesis from mmWave signal. We now proceed to introduce the design modules as follows.

- **Cross-domain Diffusion.** We propose a cross-domain diffusion model to transfer mmWave cardiac activities into ECG signals. Based on multiple iterations denoising inference of our model, the ECG synthesis is still robust even during irregular heartbeats. To achieve the domain transformation, the diffusion model consists of the following components. (i) Multichannel patchify, which is a CNN encoder to fuse multichannel mmWave data from various chest reflection points, and thus augment the mechanical activities. (ii) For the fusion condition of cross-domain diffusion, we superpose multichannel mmWave patches and Gaussian noise for the synthesis initialization at the first iteration. In the following steps, the output ECG from last iteration will take place of Gaussian noise to iterate and revise itself. (iii) Hierarchical

transformer is the core component to perform denoising operation, which takes the fusion condition as input and updates the synthesis ECG.

- **ECG Calibration Guidance.** To control the randomness and guide the denoising synthesis process. We utilize a historical ECG segment of the individual from reference device, which is introduced to constrain denoising procedure in each iteration. In particular, the calibration ECG segment is firstly processed with ECG embedding, which extracts the calibration features and fits them into tokens for transformer. In the next step, a Cross-Attention based guidance is designed to integrate calibration ECG embedding with the hierarchical transformer in Cross-domain diffusion, so as to guide the denoising process and achieve faithful synthesis. In the practical usage, any historical ECG of the individual can be directly utilized for calibration, which means *AirECG* has no requirement for model fine-tuning or frequent reference ECG measurement.

## 4 System Design

In this section, we introduce the mmWave cardiac activities extraction of *AirECG* (Sec. 4.1), the cross-domain diffusion model which transfers mechanical domain mmWave cardiac signal into ECG in electrical domain (Sec. 4.2), and *AirECG*'s ECG calibration guidance module to guide domain transformation according to historical ECG. (Sec. 4.3).

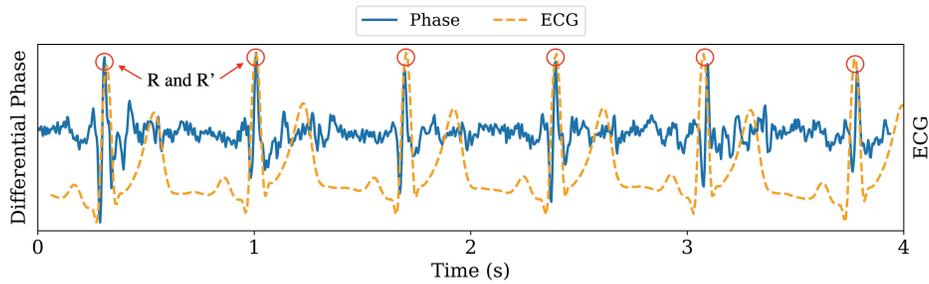


Fig. 5. The mmWave signal after cardiac activities extraction, compared with ground truth ECG

### 4.1 mmWave Cardiac Activities Extraction

For the first step of contactless ECG monitoring, we extract the mechanical cardiac activities using mmWave radar. We firstly locate the thorax wall vibration through round trip length (RTL) profile search, which identifies it by its highest vibration energy. After that, thorax wall vibration can be described as the phase signal  $\phi_t$  in Equation. 1, which is extracted from raw mmWave data after RTL profile search. However,  $\phi_t$  describes the whole chest vibration, which is a mixture of respiration, heartbeat and other motions artifacts. The chest displacement of heartbeat is in 0.2-0.5mm, while that of respiration ranges from 4 to 12mm, which is significantly larger than the displacement caused by cardiac activities [16]. To filter out irrelevant noise other than cardiac activities, the acceleration filter is designed to eliminate slow and steady motion with low acceleration, like respiration, whereas cardiac activities are rapid contraction of the myocardium, causing significant acceleration during heartbeats [49]. The acceleration can be represented as the 2nd derivative of displacement, thus the acceleration filter is defined as Equation. 2 to compute the 2nd derivative of  $\phi_t$ .

$$c(t) = \frac{(\phi_{t-3} + \phi_{t+3}) + 2(\phi_{t-2} + \phi_{t+2}) - (\phi_{t-1} + \phi_{t+1}) - 4\phi_t}{16h^2}, \quad (2)$$

Fig. 5 shows the effectiveness of acceleration filter, where the blue line is the mmWave phase after acceleration filter, the synchronous ground truth ECG is plotted in dashed yellow line. The cardiac cycle of mmWave and ECG

is the same, illustrating an obvious R' corresponds to R peak. Compared with simple band pass filter as Fig. 2(b) plotted, the phase after acceleration filter exhibits distinct cardiac cycle, thus it is employed as the mechanical domain input for our cross-domain diffusion.

## 4.2 Cross-domain Diffusion

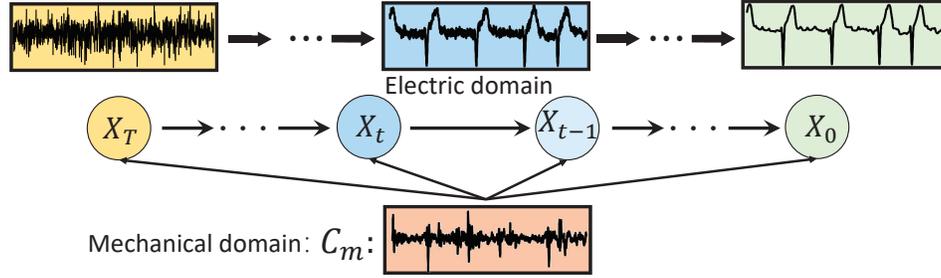


Fig. 6. The multiple iteration inference by Cross-domain diffusion.

In this subsection, we introduce cross-domain diffusion, an ECG synthesis model for mmWave-based disease monitoring. As illustrated in Fig. 6, our model takes mmWave cardiac signal and Gaussian noise as inputs, synthesizing ECG through multiple iteration denoising inference.  $X_T$  is the initial Gaussian noise, which is performed with step-by-step denoising inference into  $X_t, X_{t-1}, \dots, X_0$ , where  $t$  represents the index of denoising step,  $T$  is the number of total denoising steps, and  $X_0$  is the final ECG output. In each step of denoising inference, the mmWave cardiac signal is introduced as the condition noted as  $C_m$ . Based on such mechanical domain condition, the Gaussian noise can be gradually transformed into ECG in electrical domain. In this way, the domain transformation diffusion can be achieved.

For the cross-domain diffusion model, the core design is how to realize appropriate denoising step based on conditional domain. To this end, we design a noise predictor as shown in Fig. 8, which takes mmWave condition  $C_m$ , step index  $t$  and current intermediate ECG  $X_t$  to predict noise in the next step  $\epsilon_\theta$  and obtain ECG in next step  $X_{t-1}$ . In the following paragraphs, we will introduce how to perform forward noise diffusion and reverse denoise process respectively, so as to train the noise predictor and realize denoising step. In addition, the network architecture of noise predictor will be introduced in the section of Multichannel Patchify and Hierarchical Transformer.

**Forward noise diffusion process.** The forward noise diffusion process is proposed to add noise into the original ECG and disrupt the data distribution. In particular, Fig. 7 illustrates the steps of forward noise diffusion, which generates Gaussian noise  $N_{t-1}$  and adds it on the ECG of last step  $X_{t-1}$ , resulting in destroyed ECG  $X_t$ . Such noise diffusion step eliminates the original data distribution of ECG signal, and at the same time, generates training sample  $X_t$  for the noise predictor. Specifically, the destroyed ECG  $X_t$ , mmWave condition  $C_m$  and step index  $t$  are the inputs for the noise predictor.

We define the noise adding process as a markov process, *i.e.*,  $X_t$  is only based on  $X_{t-1}$  for adding noise. During the process, larger  $t$  represents that  $X_t$  is closer to pure Gaussian noise. To achieve this goal, we make larger noise scale as the step increases, *i.e.*, a scale parameter  $\alpha_t = 1 - 2 \times 10^{-5}t$  is defined to control the noise scale. In this way, the noise adding process of ECG can be formulated as Equation. 3.

$$\begin{aligned} X_t &= \sqrt{\alpha_t}X_{t-1} + \sqrt{1 - \alpha_t}z_1 \\ &= \sqrt{\alpha_t}(\sqrt{\alpha_{t-1}}X_{t-2} + \sqrt{1 - \alpha_{t-1}}z_2) + \sqrt{1 - \alpha_t}z_1 \end{aligned} \quad (3)$$

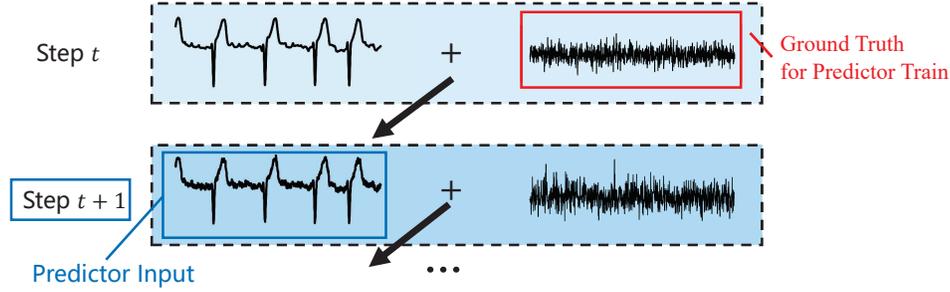


Fig. 7. The forward noise diffusion steps for predictor train.

It is noted that  $z_1, z_2$  are sampled from unit Gaussian distribution, and their shapes are the same with  $X$ , which can be marked as  $z_1, z_2 \sim \mathcal{N}(0, \mathbf{I})$ . According to the noise diffusion process in Equation. 3,  $X_{t-1}$  is required for  $X_t$ , and that implies we need to serially get the train data for noise predictor, *i.e.*, for train data in step  $t$ , we need to compute  $X_{t-1}, X_{t-2}, \dots, X_1$  serially, which has high time complexity and is hard to train. Instead, we perform further decomposition in Equation. 3, so as to achieve direct computation of  $X_t$  from  $X_0$ , as shown in Equation. 4:

$$X_t = \sqrt{\bar{\alpha}_t} X_0 + \sqrt{1 - \bar{\alpha}_t} \bar{z}_t, \quad (4)$$

where  $\bar{\alpha}_t = \prod_{i=1}^t \alpha_i$ ,  $\bar{z}_t \sim \mathcal{N}(0, \mathbf{I})$ . Based on Equation. 4, the disrupted ECG at each step can be directly computed from the pure ECG  $X_0$ , so that we can randomly select a step  $t$ , and get the disrupted ECG  $X_t$  at that step for noise predictor training. To sum up, The forward noise diffusion process adds noise to original ECG  $X_0$ , so as to get disrupted ECG  $X_t$ , which can be represented as  $q(X_t|X_0) \sim \mathcal{N}(\sqrt{\bar{\alpha}_t} X_0, \sqrt{1 - \bar{\alpha}_t} \mathbf{I})$ , where  $q$  represents forward noise diffusion process.

**Reverse denoise process.** The forward noise diffusion process  $q(X_t|X_0)$  can generate disrupted ECG  $X_t$  for noise predictor training, while the reverse denoise process is to perform noise prediction and denoising process, so as to synthesize contactless ECG. The reverse denoise process is noted as  $p_\theta(X_{t-1}|X_t)$ , where  $\theta$  represents neural network parameters that can be trained. The objective of denoising process is to make synthesized ECG distribution  $p_\theta(X_0)$  close to original ground truth ECG  $q(X_0)$ , which can be achieved through minimizing their Kullback-Leibler (KL) divergence as proved in prior works [23, 37]:

$$\begin{aligned} \theta &= \arg \min_{\theta} D_{KL}(q(X_0) || p_\theta(X_0)) \\ &= \arg \min_{\theta} D_{KL}(q(X_{t-1}|X_t, X_0) || p_\theta(X_{t-1}|X_t, C_m)), \end{aligned} \quad (5)$$

where  $q(X_{t-1}|X_t, X_0)$  represents the ground truth denoise process to get denoised ECG  $X_{t-1}$  based on intermediate ECG  $X_t$  and pure ECG  $X_0$ , while  $p_\theta(X_{t-1}|X_t, C_m)$  denotes the denoise process through the noise predictor based on mechanical domain mmWave condition  $C_m$ . In particular, the ground truth denoise process to get  $X_{t-1}$  is a Gaussian process, noted as  $q(X_{t-1}|X_t, X_0) \sim \mathcal{N}(\mu_{t-1}, \sigma_t^2 \mathbf{I})$ , where  $\sigma_t = \frac{(1 - \bar{\alpha}_{t-1})(1 - \alpha_t)}{1 - \bar{\alpha}_t}$ . However, such Gaussian process of  $q(X_{t-1}|X_t, X_0)$  can not be used for neural network training directly since it is not suitable for gradient descent. Instead, our goal is to get the Gaussian mean  $\mu_{t-1}$  from neural network output, noted as  $\mu_\theta$ , which can be utilized to sample denoised ECG  $X_{t-1}$ . The ground truth  $\mu_{t-1}$  can be computed as:

$$\mu_{t-1} = \frac{\sqrt{\bar{\alpha}_{t-1}}(1 - \alpha_t)X_0 + \sqrt{\alpha_t}(1 - \bar{\alpha}_{t-1})X_t}{1 - \bar{\alpha}_t}, \quad (6)$$

thus we can perform network training through minimizing the mean square error (MSE) between mean of ground truth  $\mu_{t-1}$  and predicted mean  $\mu_\theta(X_t, t, C_m)$ , where  $C_m$  is mmWave mechanical domain activities as a condition.

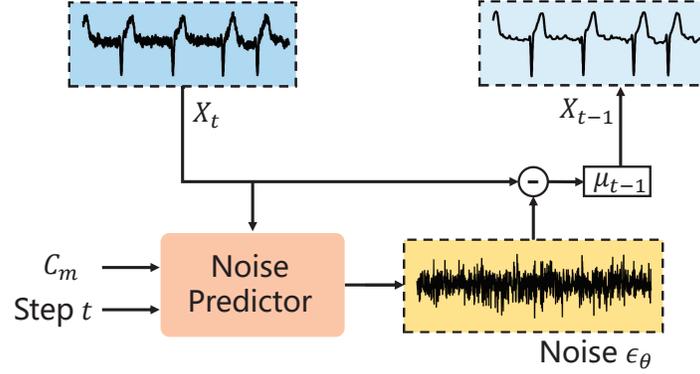


Fig. 8. A denoising step: predict noise in ECG  $X_t$  and synthesize  $X_{t-1}$ .

The training objective in Equation. 5 can be updated as:

$$\theta = \arg \min_{\theta} \|\mu_{t-1} - \mu_{\theta}(X_t, t, C_m)\|^2, \quad (7)$$

which is the core of noise predictor training. It should be noted that  $\mu_{\theta}$  is not directly output by the noise predictor, in particular, the predictor predicts noise  $\epsilon_{\theta}(X_t, t, C_m)$  in  $X_t$ , and  $\mu_{\theta}$  can be computed from  $\epsilon_{\theta}$  as the following Equation. 8 shows:

$$\mu_{\theta}(X_t, t, C_m) = \mu_{\theta}(X_t, t, \epsilon_{\theta}) = \frac{1}{\sqrt{\alpha_t}} \left( X_t - \frac{1 - \alpha_t}{\sqrt{1 - \bar{\alpha}_t}} \epsilon_{\theta}(X_t, t, C_m) \right), \quad (8)$$

therefore, to get  $\mu_{\theta}$ , the noise predictor output  $\epsilon_{\theta}$  is the only variable required for training. After the training of noise predictor  $\epsilon_{\theta}$ , we can achieve the reverse denoise inference based on Gaussian sampling:

$$\hat{X}_{t-1} = \mu_{\theta}(\hat{X}_t, t, \epsilon_{\theta}) + \sigma_t z \quad (9)$$

where  $z \sim \mathcal{N}(0, \mathbf{I})$ ,  $\epsilon_{\theta}(\hat{X}_t, t, C_m)$  is the noise prediction output by the predictor. Besides, when  $t = 1$ , set  $\sigma_t = 0$  for the final  $X_0$  inference. To sum up, the reverse denoise process supports multiple iteration inference using a noise predictor, which is trained through the optimization of Equation. 7, and the inference process is based on Equation. 9.

Algorithm. 1 is the pseudo-code of cross-domain diffusion training process, where we take mechanical domain mmWave and electrical domain ECG ground truth as input. During the training process, a step  $t$  is randomly selected as a sample, where forward noise diffusion process (Equation. 4) is executed to get disrupted ECG  $X_t$ . Then we perform reverse denoise process to optimize the noise predictor  $\epsilon_{\theta}$  through minimizing the mean square error in Equation. 7. For the multiple iteration inference of cross-domain diffusion in Algorithm. 2, mechanical domain mmWave is input into the noise predictor to predict the noise  $\epsilon_{\theta}$  in  $\hat{X}_T, \dots, \hat{X}_1$  sequentially. In each iteration, the  $\hat{X}_{t-1}$  can be synthesized through Gaussian sample in Equation. 9, thus the initially noisy data  $X_T$  can be transformed into electrical domain ECG step-by-step.

**Multichannel Patchify and fusion condition.** The reverse denoising process requires a noise predictor  $\epsilon_{\theta}$ , which can take the input not only from denoising status (step  $t$  and intermediate ECG  $X_t$ ), but also mmWave cardiac activities as condition  $C_m$ . To achieve the goal, two main challenges with the mmWave issues need to be addressed, the first one is how to extract distinct features from the multichannel mmWave data, which can provide precise cardiac information through multiple antennas on mmWave radar. The second challenge is how to fuse intermediate ECG  $X_t$  with mmWave condition  $C_m$ , so as to be suitable for multiple iteration denoising inference.

**Algorithm 1** Cross-domain Diffusion Training

---

**Input:**  $C_m$ : mmWave condition;  
 $X_0$ : ECG ground truth;  
**Output:** Noise Predictor  $\epsilon_\theta$

- 1: Set total diffusion steps  $T$
- 2: Get Parameters  $\alpha_t$  and  $\bar{\alpha}_t$
- 3: **while**  $\epsilon_\theta$  not converged **do**
- 4:   Randomly sample  $t$  in  $[1, T]$
- 5:   Sample  $z \sim \mathcal{N}(0, \mathbf{I})$
- 6:   Get predictor input  $X_t(X_0, t)$
- 7:   Get ground truth mean  $\mu_{t-1}(X_t, X_0)$
- 8:   Predict Noise  $\epsilon_\theta(X_t, t, C_m)$
- 9:   Get  $\mu_\theta(X_t, t, \epsilon_\theta)$
- 10:   Minimize  $\|\mu_{t-1} - \mu_\theta(X_t, t, \epsilon_\theta)\|^2$
- 11: **end while**

---

**Algorithm 2** Cross-domain Diffusion Inference

---

**Input:**  $C_m$ : mmWave condition;  
 $\epsilon_\theta$ : Noise Predictor;  
**Output:** Synthesized ECG  $\hat{X}_0$

- 1: Set total diffusion steps  $T$
- 2: Get Parameters  $\alpha_t$ ,  $\bar{\alpha}_t$  and  $\sigma_t$
- 3: Sample  $z \sim \mathcal{N}(0, \mathbf{I})$
- 4: Gaussian initialization  $\hat{X}_T = \sigma_t z$
- 5: **for**  $t = T, \dots, 1$  **do**
- 6:   Predict Noise  $\epsilon_\theta(X_t, t, C_m)$
- 7:   Get  $\mu_\theta(X_t, t, \epsilon_\theta)$
- 8:   Sample  $z \sim \mathcal{N}(0, \mathbf{I})$
- 9:   Sample  $\hat{X}_{t-1} = \mu_\theta + \sigma_t z$
- 10: **end for**
- 11: return  $\hat{X}_0$

---

In terms of feature extraction in multichannel mmWave, a CNN-based multichannel patchify is proposed to split mmWave data into patches, and each patch can be considered as a high-dimension token with cardiac features as shown in Fig. 9. In particular, the original form of mmWave is multichannel sequential data, in the 1D-shape of  $L \times C$ , where  $L$  is the length of mmWave and  $C$  is the number of channels. However, 1D-shape does not suit the denoising process, since the denoising is directly performed on the whole segment, but not sequentially over time. Hence we adjust the mmWave into 2D-shape of  $I \times I \times C$ , where  $I = \sqrt{L}$ . In addition, to extract cardiac features from multichannel data, we introduce a CNN framework, which computes the convolutions of each channel and performs summation on all channels. Such a framework can achieve the channel fusion function similar to beamforming in [20], so as to boost cardiac information through multiple channels, since more channels can provide higher resolution for mmWave radar. To extract distinct multichannel features, we custom-designed the convolution layer as it can split mmWave patches in the shape of  $p \times p$  and transfer them into mmWave tokens for the hierarchical transformer in the next part. Specifically, the kernel size and stride step of convolution are the same as patch size  $p = 2$ , and the number of feature maps is set to  $d = 384$ . Thus the multichannel mmWave data can be patchified into  $T$  tokens with  $d$  dimensions, where  $T = (I/p)^2$ . Those tokens are processed with standard 2D positional encoding to fit transformer framework [18], and the tokens are noted as mmWave condition  $C_m$ , which can be fused with intermediate ECG  $X_t$ .

The condition information controls the multiple iteration denoising of diffusion model, which can be class labels (e.g., dog or cat) [31] and prompt text [34] in prior image synthesis works. However, different from the condition information in prior works, the mmWave condition is strongly correlated with synthesized ECG in time domain since they are synchronous. To maintain such relationship during denoising process, we design a fusion condition which adds the tokens of mmWave and intermediate ECG together. Furthermore, mmWave condition and the intermediate ECG also share the same 2D positional encoding to maintain synchronization in time domain. In this way, the mmWave mechanical domain cardiac activities can be deeply fused with intermediate ECG, thus the domain transformation can be achieved during each iteration of ECG denoising process.

**Hierarchical Transformer.** Our cross-domain diffusion is a hybrid architecture neural network, where the fusion condition is convolution-based, and the backbone is transformer-based. As shown in Fig. 10, the transformer backbone is hierarchical, and we define 6 hierarchical transformer blocks in our implementation, which are scalable on demand. The backbone takes fusion condition and time step  $t$  as inputs, and estimates the

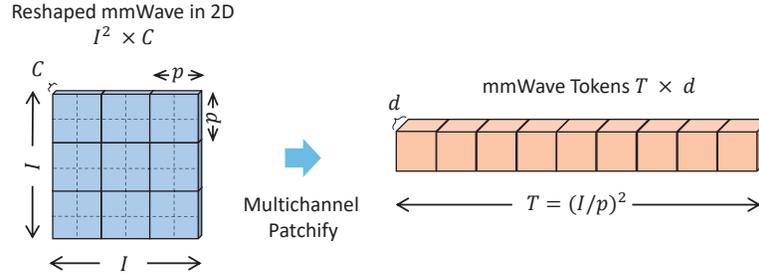


Fig. 9. Multichannel patchify: convolution fusion of multichannel mmWave data into tokens.

noise level  $\epsilon_\theta$  in intermediate ECG  $X_t$ . In the following paragraphs, we will introduce the modules in transformer backbone from bottom to the top.

- Time step Embedding is to transform time step  $t$  into vectors, which are in the same dimension of transformer backbone ( $d = 384$ ). Since earlier denoising steps only require dealing with some rough features, while detailed denoising is needed in latter steps. Thus it is necessary to combine time step information into the noise predictor. A sinusoidal timestep embedding is implemented with 2 linear layers and SiLU activation.
- Adaptive Layer Normalization (adaLN) is a design to incorporate the time step deeply into neural network. Prior works have proved the effectiveness of adaLN in image synthesis [17, 31], thus we design adaLN modules after each layer normalization operation to deeply augment time step embedding. In the transformer block, an adaptive scale operation is also performed after Self-Attention and Feedforward.
- Multi-Head Self-Attention is the core module to transfer mechanical domain cardiac activities into electrical domain. It can capture auto correlation information such as atrial and ventricular activities and extract high-level representations implicit both in mmWave and intermediate ECG. The hybrid architecture of self attention and convolution in multichannel patchify can improve the training efficiency, since their advantages of local feature extraction and global receptive field can be cooperated.

### 4.3 ECG Calibration Guidance

The cross-domain diffusion is a type of generative model, which shows randomness nature due to Gaussian initialization and denoising. To avoid the negative impacts of randomness on synthesized ECG fidelity, the ECG calibration guidance is introduced to guide and control the generative process, so as to support faithful and high accuracy contactless ECG.

According to the multiple iteration denoising inference mentioned above, the randomness is unavoidable due to the generative ECG synthesis. Despite the fact that mmWave condition  $C_m$  has controlled the randomness to some extent, further guidance is still worthy of user's personalized information, especially for new users who are not involved in the train datasets. A simple idea is to fine-tune a personal model with new user's data, while it requires synchronized ground truth data and large computing resources, thus it is not applicable in practical scenarios. To achieve personal guidance without the requirement of synchronous data and fine-tuning, we design ECG calibration guidance as a plug-in to calibrate synthesized output. Such calibration can be achieved through historical ECG ground truth from specific user. We design a Multi-Head Cross-Attention based guidance to calibrate the cross-domain diffusion model. The Cross-Attention can learn the potential correlation between synthesized ECG and historical ECG, thus controlling the denoising step to fit the personal features in calibration ECG.

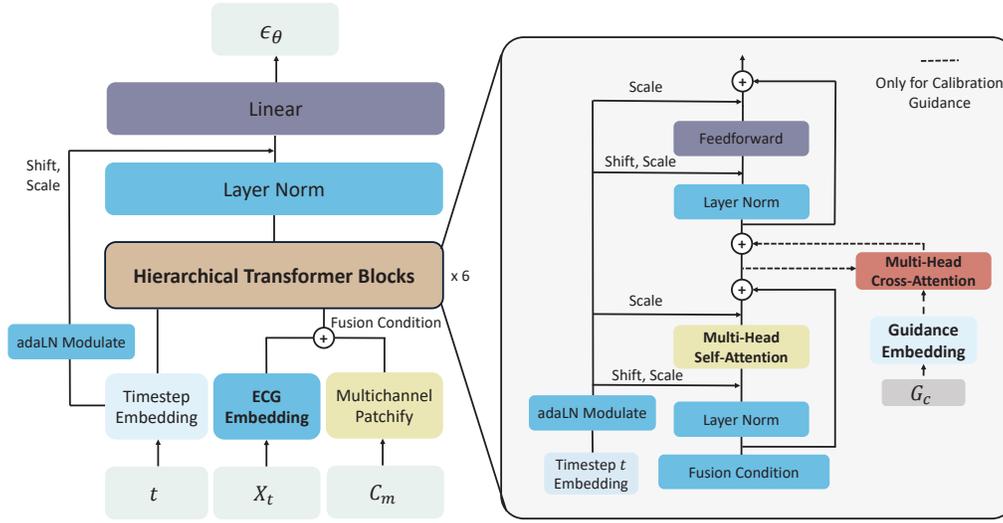


Fig. 10. The noise predictor architecture of Cross-domain Diffusion.

The work flow of ECG calibration guidance is illustrated as the dashed line in Figure. 10. In particular, the calibration ECG for guidance is noted as  $G_c$ , which is firstly fed into Guidance embedding, so as to achieve ECG features extraction and map them into the dimension  $d$  of transformer backbone. Then the latent features of synthesized ECG are introduced from the self-attention layer, which is further mapped into query vector  $Q_X$  in our Cross-Attention. At the same time, the guidance embedding is mapped into key  $K_G$  and value  $V_G$  vectors of Cross-Attention. The trainable attention score can be computed as follows:

$$\text{CrossAttention}(Q_X, K_G, V_G) = \text{softmax}\left(\frac{Q_X K_G^T}{\sqrt{d}}\right) V_G, \quad (10)$$

where the calibration ECG  $G_c$  can be fused with synthesized ECG  $X$  for the following procedure, achieving the calibration guidance. It should be noted that the calibration guidance can be omitted if there is no historical ECG for guidance, however, the synthesized ECG fidelity may decline on unseen participants as described in Sec.6.5.

## 5 Experiment Setup

**Hardware and software toolkits.** Fig. 11(a) shows the mmWave radar and ground truth ECG patch in our experiment, and it should be noted that mmWave radar is packaged into a 3D printed shell. The IWR1443BOOST mmWave radar [8] and DCA1000EVM [7] from Texas Instrument are deployed as the mmWave front end to capture mechanical domain cardiac activities. The IWR1443 radar is equipped with 2 Tx and 4 Rx to collect multichannel mmWave data, which form 8 virtual channels in total. The radar emits 250 frames per second, with each frame containing 2 chirps from each Tx. Each chirp is a continuous wave that lasts 65ms and consists of 512 sample points. The mmWave data is transmitted from the DCA1000EVM to our laptop through Ethernet, where we develop a packet sniffing program to collect data in real-time. For ground truth ECG data collection, the CarePulse [4] ECG patch is employed with a sample rate of 125Hz, which shares a timestamp with the mmWave radar via network time protocol (NTP). The ECG ground truth is analyzed by professional ECG labeling software (ECG Pro from CarePulse [1]), the software marks the category of heart beat (normal, AF, etc.) and is checked by expert technicians manually. The raw data of mmWave and ground truth ECG are transmitted to a data server for

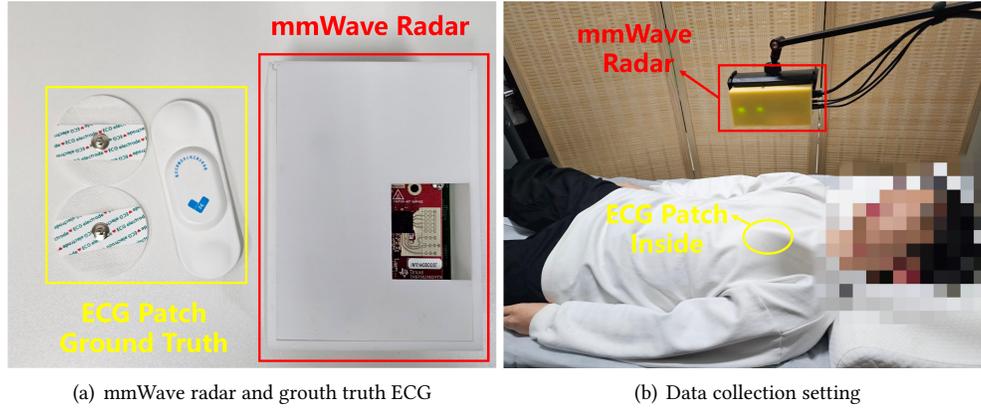


Fig. 11. Devices and data collection setting of experiment.

further signal processing using Python and libraries of NumPy, SciPy. Our deep learning model is implemented with Pytorch, which is deployed on a server with 2 NVIDIA RTX3080 GPUs and Intel Xeon E5-2696v3 CPU.

**Participant enrollment and disease diagnosis.** For the participant enrollment, this research is approved by the institutional review board (IRB) of Peking University Third Hospital (PUTH). 32 participants (21 males, 11 females, 47.6 years old on average) are enrolled from May to December 2023. Among all participants, 16 were inpatients from cardiology ward, while the remaining 16 participants were enrolled as healthy control group. The participants of cardiac disease group may be diagnosed with one or more diseases such as AF, heart failure, *etc.* In terms of AF patients, both persistent and paroxysmal AF patients were recruited in our experiment to increase the diversity. It should be noted that all the participants have been informed of the experiment content and provided informed consent in advance, which is an agreement to share the data for research. The experiments were performed in the office room and cardiology ward of the hospital. The ground truth ECG data from patch devices were annotated by three expert technicians to mark heartbeats into rhythm labels, including normal sinus rhythm (NSR), atrial fibrillation (AF), and other 5 common arrhythmia types: premature ventricular contraction (PVC), atrioventricular block (AVB), premature atrial contraction (PAC), tachycardia and bradycardia. In total, 109,598 heartbeats are captured by mmWave and included in our datasets.

**Data collection setting.** The data collection setting of our experiment is exhibited as Fig. 11(b), where a participant is lying under the mmWave radar and the distance is about 20~60cm away from antennas. The participant also wears an ECG patch under his/her clothes, which collects II-lead ECG as ground truth. The data collection process lasts 1 continuous hour for each participant. In total, the duration of synchronous mmWave and ECG is 32 hours, and we split data into 5s length segments for training and evaluation. Specifically, we resample the original sample points (mmWave: 1250, ECG: 625 points) into 1024 (32×32) points to fit the multichannel patchify of cross-domain diffusion. In addition, to augment training datasets, we shift segments by 1 second for data augmentation, which can improve model performance and avoid over-fitting.

## 6 System-level Evaluation

In this section, we first evaluate the performance of *AirECG* compared with state-of-the-art contactless ECG monitoring methods on normal participants and participants with diseases, separately. Furthermore, we validate the effectiveness of our system on user-independent datasets, *i.e.* new participants apart from train datasets. Then we show *AirECG*'s performance under 3 types of cardiac disease monitoring, *i.e.*, AF, PVC, and AVB,

respectively. Additionally, to demonstrate the ability of contactless ECG for downstream applications, we perform arrhythmia detection as a further experiment. In the end, we perform micro-benchmark experiments to validate the effectiveness of *AirECG*'s modules and parameter configurations individually.

## 6.1 Overall performance

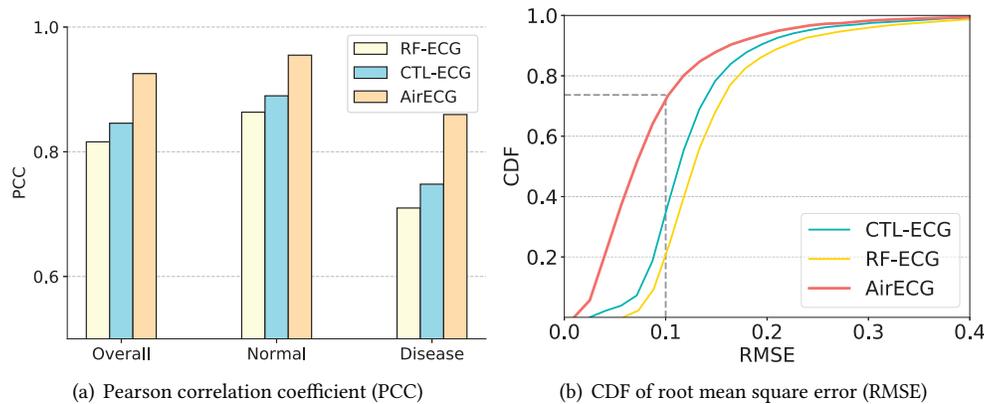


Fig. 12. Signal fidelity between synthesized ECG and ground truth.

**Evaluation metrics.** To train and evaluate *AirECG*, we perform the 5-fold cross-validation based on user-dependent and user-independent (in the following Sec. 6.2) principle respectively. For user-dependent datasets, we randomly split the overall datasets into 5 folds, with each fold containing 20% of the data. To evaluate our model performance, we train *AirECG* 5 times in total, and one fold is selected as the test datasets each time, while the other folds are utilized to train the model (70% train, 10% validation). The test dataset's results for each fold are summarized as the final evaluation performance. In particular, we utilize the following 5 metrics to evaluate the quality of synthesized ECG signal.

- *Pearson correlation coefficient (PCC)*, which is a statistical measure of the amplitude and trend similarity between the synthesized ECG and ground truth. The value of PCC ranges from -1 to 1, and higher value implies synthesized ECG has higher correlation with ground truth. The relationship is considered to be strong correlation when PCC is larger than 0.8.
- *Root mean square error (RMSE)*, which stands for the amplitude error between synthesized ECG and ground truth. We calculate root of mean square error among all sample points of normalized ECG, so as to represent the amplitude error of the entire segment.
- *Qualified monitoring rate*. To evaluate the cardiac events in ECG, we employ Neurokit2 [27] to identify ECG and corresponding feature points, which is an open source software for ECG signal processing. While low quality ECG may not be successfully identified due to the limitation of synthesis model and ambient noise, thus we utilize the identified rate as a metric for synthesis quality.
- *R-R interval error*, which is the interval error between synthesized ECG and ground truth. As the primary cardiac event corresponding to heart systole, the R-R interval is related to heart rate variability and heart disease like AF. Once the R peaks in ECG are identified by Neurokit2, R-R intervals can be computed in milliseconds.
- *T-wave timing error*, which is the timing error between T-waves in the synthesized ECG and ground truth. T-wave represents the potential changes of ventricular repolarization and is related to ventricular diastole. Similar to R-R interval, T-wave can be identified for timing error computation in milliseconds.

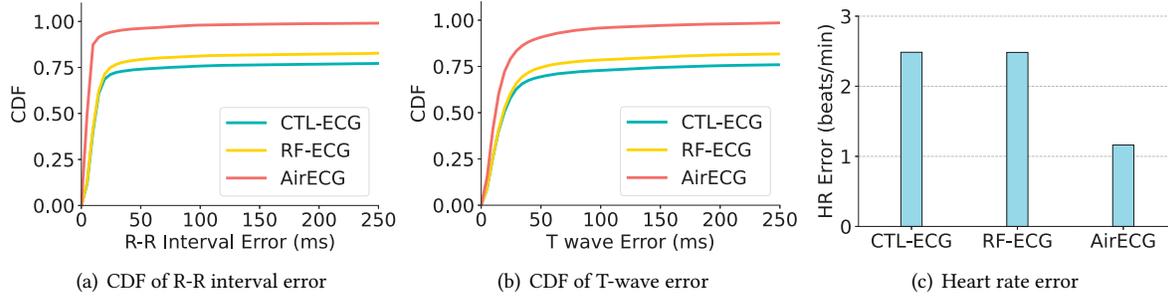


Fig. 13. Timing error of cardiac events. (a) is the Cumulative Distribution Function (CDF) plot of R-R interval error, (b) is the CDF plot of T-wave timing error. Some heartbeats fail to be identified in baselines, but they are valid in *AirECG*, thus we set them to an error of 1s for baselines. (c) is the heart rate error among all valid segments of three algorithms.

**Comparison baselines.** We implement two state-of-the-art contactless ECG baselines for performance comparison, the first baseline is RF-ECG [42], which utilizes a conditional GAN based model to synthesize ECG from wireless signal. The cGAN of RF-ECG takes wireless signal as condition and generates ECG based on a convolution encoder-decoder framework, noting that the cGAN model outputs generated ECG through one step inference. The second baseline is a contactless ECG monitoring method with mmWave radar [11], known as CTL-ECG. Such method is based on a sequence to sequence (Seq2Seq) architecture, which extracts cardiac features using a temporal-spatial decoder, while synthesizing ECG from cardiac features using auto-regressive CNN. The auto-regressive principle implies that the model outputs one sample point per inference, that is, for 1024 sample points in 5s segment, 1024 iteration inference is required. These points are generated sequentially, meaning that points in the front can not be revised after inference. In contrast to previous approaches, *AirECG* performs multiple iteration denoising inference, which can synthesize ECG step-by-step, thus the correlation between sample points can be included in revising the output ECG. In this way, *AirECG* can achieve accurate ECG monitoring even during abnormal heartbeats.

**Accuracy of ECG morphology.** To evaluate ECG morphology in signal amplitude and trend, the PCC and RMSE are employed as metrics. Referring to Fig. 12(a), we can observe that, (i) for the overall datasets, *AirECG* can achieve the highest PCC at 0.925 with 7.9%~13.4% improvement, which implies the effectiveness of cross-domain diffusion from mechanical domain mmWave to electrical domain ECG. Besides, CTL-ECG (0.846) shows better performance than RF-ECG (0.816), demonstrating the advantage of sequentially generation than one step inference. (ii) We separately evaluate the PCC of health participants (Normal) and participants with cardiac diseases (Disease). It can be found that for health participants, the synthesized signals have a strong correlation (PCC higher than 0.8) with ground truth, when using any of three methods: *AirECG* (0.95), RF-ECG (0.864) or CTL-ECG (0.890). Such performance owes to the regular heartbeats of healthy participants, which are uncomplicated to perform domain transformation. (iii) However, RF-ECG and CTL-ECG fall short during disease monitoring, showing PCC of 0.710 and 0.748 respectively. The performance decline of baselines is caused by the model limitation for abnormal ECG synthesis, which demonstrates irregular rhythm and morphology. *AirECG* overcomes this problem with the help of multiple iteration denoising inference, showing a strong correlation PCC of 0.860 (15.0%~21.1% improvement).

For the RMSE evaluation between synthesized ECG and ground truth, Fig. 12(b) plots the cumulative distribution function (CDF) of *AirECG* and baselines. We can find that *AirECG* outperforms others in all error ranges. In particular, 73.67% of BreathAnalyzer’s measurements are within a normalized RMSE of 0.1, compared with that of 23.76% and 37.55% for RF-ECG and CTL-ECG, respectively. The smaller of RMSE implies better performance on

Table 1. Performance on user-independent datasets.

Model	PCC(N)	PCC(D)	RMSE(N)	RMSE(D)
RF-ECG	0.755	0.654	0.223	0.374
CTL-ECG	0.754	0.643	0.224	0.378
<b>Ours</b>	<b>0.853</b>	<b>0.819</b>	<b>0.182</b>	<b>0.290</b>

Table 2. PCC under various diseases.

Model	PVC	AVB	AF
RF-ECG	0.908	0.809	0.675
CTL-ECG	0.898	0.770	0.620
<b>Ours</b>	<b>0.959</b>	<b>0.911</b>	<b>0.802</b>

the amplitude of synthesized ECG, which helps to monitor diseases such as myocardial hypertrophy or myocardial necrosis.

**Accuracy of cardiac events.** The accuracy of cardiac events (R-R interval, T-wave) is of significant importance for disease monitoring, as indicators. However, contactless ECG may synthesize unqualified ECG signals, lacking QRS-complex or T-wave, due to user's body movement and ambient noise, therefore, the qualified monitoring rate is employed to evaluate the quality of ECG wave. The open source ECG processing tool Neurokit2 is used to identify the QRS-complex and T-wave in ECG, while we find that the qualified monitoring rates for RF-ECG, CTL-ECG and *AirECG* are 67.69%, 72.81% and 86.95% respectively. We summarize the reasons as following, (i) As previously mentioned, the abnormal heartbeats bring challenge for ECG synthesis, causing more unqualified ECG. While *AirECG* shows better performance in qualified monitoring rate, this exhibits the effectiveness of ECG synthesis through multiple iteration denoise. (ii) The body movements and ambient noise are more serious and unavoidable in our datasets, since data collection is in shared wards of multiple patients (personal activities cause unavoidable ambient noise), and patients may have more body movement compared with normal subjects due to physical discomfort. (iii) Neurokit2 has limited ability to process ECG with noise, which reduces the qualified monitoring rate level.

Furthermore, we evaluate the timing accuracy of contactless ECG, as the ECG wave is successfully identified. Fig. 13(a) is a CDF plot of R-R interval error of RF-ECG, CTF-ECG and *AirECG*. Due to the fact that partial R peaks can be identified by *AirECG* while not validated in RF-ECG and CTF-ECG, we set the timing error of them into 1s (corresponding to typical 60 beats/min heart rate). In this way, the average errors of them are 228.3ms, 171.5ms and 10.3ms respectively. Our method outperforms others in all error ranges of the CDF plot in Fig. 13(a). We consider that the limitations of prior works are mainly caused by low quality ECG during disease monitoring. Especially for the main disease in our datasets: atrial fibrillation (AF), the heartbeats are completely irregular, resulting in high errors in prior ECG synthesis.

Fig. 13(b) plots the CDF of T-wave timing error, where the false identified T-wave in RF-ECG and CTF-ECG is also set to 1s. The mean errors of RF-ECG, CTF-ECG and *AirECG* are 251.3ms, 192.6ms and 24.5ms respectively. Compared with R-R interval, the errors of T-wave increase for all three algorithms, since the cardiac activities of T-wave (ventricular repolarization) are weaker than that of R peak (ventricular depolarization). Therefore, it is harder to synthesize T-wave, introducing higher error.

In the next step, we employ the effective heartbeats identified by Neurokit2 from RF-ECG, CTF-ECG and *AirECG* respectively, so as to estimate the heart rate error of the effective heartbeats. Among the heartbeats of CTL-ECG, the average heart rate error is 2.49 beats/min, while that of RF-ECG is 2.48 beats/min. It implies that for the effective segments, the performance between those two systems is similar. It is noted that CTL-ECG has more effective segments (72.81%) than RF-ECG (67.69%) as we mentioned before. The average heart rate error of *AirECG* is 1.16 beats/min, which shows significant improvement with the help of cross-domain diffusion model.

## 6.2 User-independent performance

To evaluate the generalization of *AirECG*, we implement a user-independent 5-fold cross-validation experiment, *i.e.*, the test participants (6 subjects in each fold including normal and diseased condition) have not been included

in the model training datasets (the other 26 subjects). Table 1 shows the PCC and RMSE on the test datasets, which are separated into normal subjects (N) and subjects with disease (D). We can observe that,

(i) For our algorithm the average PCC of normal subjects (0.853) and subjects with disease (0.819) both show strong correlations with ground truth ECG, while the PCCs of baselines decline to lower than 0.8. The advantage of *AirECG* comes from the calibration guidance module, which incorporates the reference ECG from each participant into the ECG synthesis. Although the network parameters are not fine-tuned by the personal data from test participants, the calibration guidance module can extract ECG morphology information from reference ECG and guide the denoising synthesis process. Thus contactless ECG can be accurate and avoid the impact from randomness of generative model. (ii) For the RF-ECG and CTL-ECG baselines, the user-independent performance declines compared with evaluation in previous works. One possible reason is the limitation of our ground truth ECG device, since the ECG patch provides a non-standard lead. Compared with standard lead in prior works, ECG patch may cause more ECG morphology variation among different subjects, thus impacting the result of user-independent evaluation.

### 6.3 Performance among Typical Types of Arrhythmia

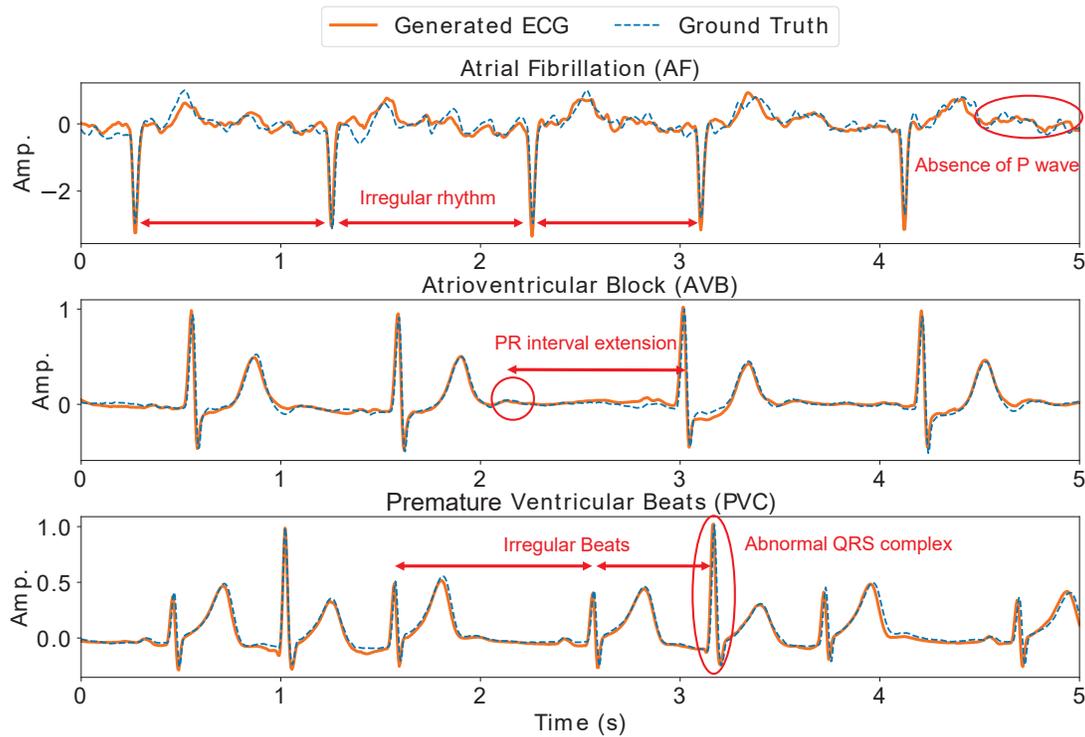


Fig. 14. Examples of *AirECG* during different diseases. (a) Atrial fibrillation (AF): P-wave disappears with irregular QRS-complex. (b) Atrioventricular block (AVB): PR interval extends as the 2-3s in figure, showing limited atrioventricular conduction. (c) Premature ventricular beats (PVC): irregular beats and abnormal QRS-complex at 1 or 3 s

We have validated *AirECG*'s performance on ECG morphology and events in previous experiments. In this subsection, we discuss the performance on specific diseases, *i.e.*, PVC, AVB and AF. We evaluate the PCC of synthesized ECG under various diseases, which is listed in Table 2. It can be summarized that, (i) *AirECG*

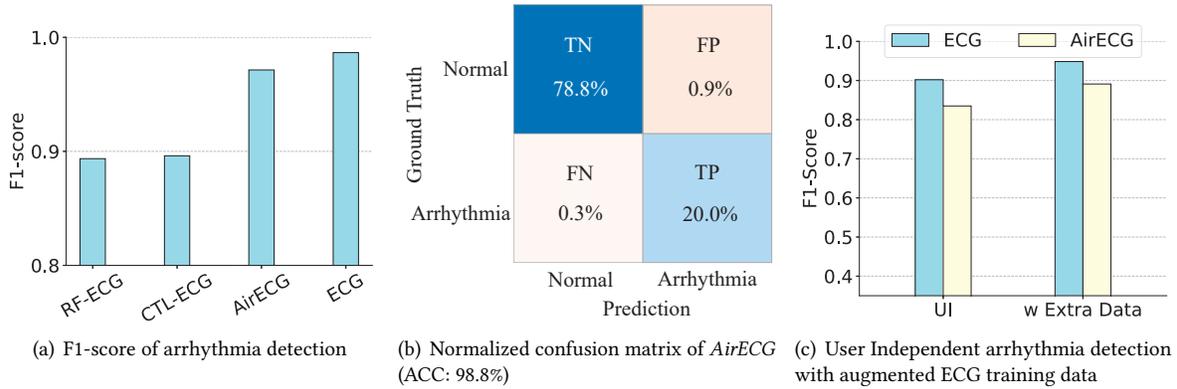


Fig. 15. Arrhythmia detection performance compared with ECG patch and mmWave generated ECG.

outperforms baselines on all diseases in our datasets, which implies that cross-domain diffusion helps synthesizing of abnormal heartbeats, through multiple iteration denoising inference. (ii) The performance under PVC or AVB is better than that of AF, and we consider such phenomenon is caused by the characteristics disease. In particular, AF shows continuously irregular heartbeats with abnormal ECG feature points (even feature point disappear), while PVC and AVB still shows integral ECG feature points with partial irregular heartbeats. (iii) We consider the main issue limits the performance of RF-ECG and CTL-ECG in the overall test datasets is the ECG fidelity decline during AF, since their PCC drops to 0.675 and 0.620 during AF occurs. While AF accounts for a large portion of our disease datasets.

Fig. 14 shows the contactless ECG and corresponding ground truth under three diseases respectively. The synthesized ECG signals present the same periodicity, amplitude and trend compared with the ground truth under AF, AVB and PVC. Slight misalignment randomly happens during ECG monitoring, which will not impact the diagnosis for diseases. For diseases diagnosis with specific ECG feature points, (i) In terms of AF, we observe the disappearance of the P-wave along with continuously irregular QRS-complex. (ii) The ECG of AVB shows an PR interval extension at the 2-3s of the segment, implying limited atrioventricular conduction. (iii) PVC shows irregular beats and abnormal QRS-complex at 1 and 3 s. It can be summarized that *AirECG* enables disease monitoring and has the ability to present the features of specific diseases.

#### 6.4 Arrhythmia Detection using Contactless ECG

The *AirECG*'s performance under different diseases has been validated in the previous subsection. Furthermore, we discuss the performance of ECG downstream tasks, *i.e.*, arrhythmia detection. We use the existing state-of-the-art ECG arrhythmia detection algorithm [22] to evaluate the ability of contactless ECG. In particular, the ECG data from patch device and artificial annotation are utilized for detector training. The arrhythmia detector in [22] has been shown to achieve cardiologist-level arrhythmia detection. For the evaluation, we first feed the test mmWave data into ECG synthesis model and get the synthesized ECG, then the ECG is fed into arrhythmia detector for the final evaluation. We also feed the ground truth ECG into the detector to assess the performance of gold standard device. For arrhythmia detection, the F1-score is a common metric used to evaluate model performance, which helps avoid the negative impact of unbalanced data distribution.

Fig. 15(a) plots the F1-score of arrhythmia detection on RF-ECG, CTL-ECG, *AirECG* and the ground truth ECG. We can observe that, (i) The ground truth ECG can achieve a detection F1-score of 0.987, demonstrating the gold standard of arrhythmia detection. While *AirECG* can achieve a similar performance of 0.9714, showing the

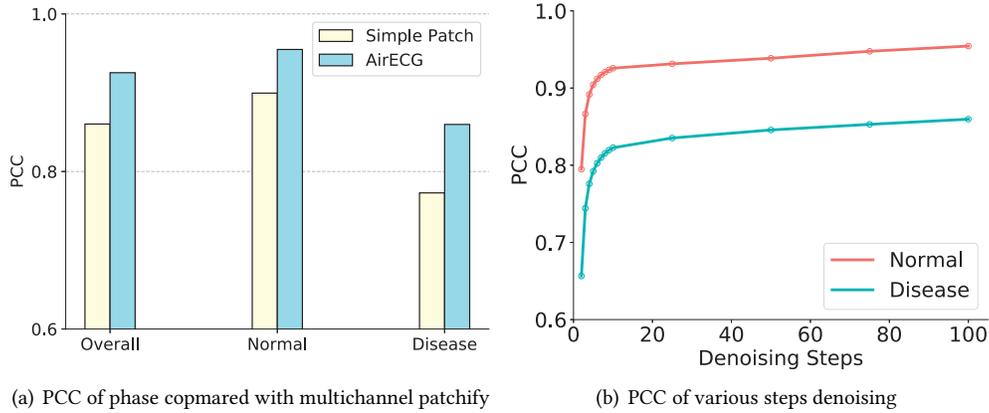


Fig. 16. Experiment for multichannel patchify and various denoising steps.

effectiveness of contactless disease monitoring. (ii) We find that *AirECG* shows great improvement compared with other contactless ECG baselines. We consider the great improvement is mainly because the performance of these baseline suffers greatly when facing disease patients, which leads to the misjudgment of arrhythmia detector. Fig. 15(b) shows the normalized confusion matrix of *AirECG*. The results show that, *AirECG* achieves 98.8% accuracy, 98.88% specificity and 98.58% recall, showing the reliability for arrhythmia detection.

To evaluate the generalization of arrhythmia detection, we conduct a user-independent validation on *AirECG* and ground truth ECG. Which means that the participants in test datasets are isolated from training datasets. Fig. 15(b) illustrates the F1-score of *AirECG* (0.8350) and ground truth ECG (0.9020) under user-independent validation conditions. The performance still maintains reasonable level for disease screening, and we consider the performance decline is caused by insufficient ECG training data and weak generalization of arrhythmia detector in [22]. To this end, we employ open source ECG data in [19, 29] for arrhythmia detector training, and the F1-score of *AirECG* and ground truth ECG increases to 0.8911 and 0.9484 respectively. We believe that the generalization of arrhythmia detection can benefit from larger quantities of training data.

## 6.5 Micro-benchmark Evaluation

**Effectiveness of multichannel patchify.** To validate the effectiveness of multichannel patchify and hybrid CNN-Transformer framework in *AirECG*, we introduce the baseline that replaces multichannel patchify with a single-channel MLP projection layer, similar to the setting in vision transformer [18]. The baseline is named Simple Patch, and Fig. 16(a) shows the performance between *AirECG* and simple patch. Overall, it can be observed that the PCC declines from 0.925 to 0.860, with the PCC of normal participants decreasing from 0.955 to 0.900. Especially for disease patients, the PCC declines from 0.860 to 0.773. We can summarize that, multichannel mmWave is essential to improve the accuracy of ECG monitoring, in particular patients with abnormal heartbeat. Multichannel data has an enhancing effect on the heartbeats that are hard for ECG measurement. Besides, the hybrid CNN-Transformer framework performs better than pure transformer, since the CNN can be used for multichannel feature extraction at front layers, and the convolution operation is more suitable for temporal mmWave data.

**Impact of denoising steps.** In previous experiments, it has been demonstrated that cross-domain diffusion benefits ECG synthesis from mmWave signals, particularly through multi iteration denoising inference. To evaluate how the step count of denoising impacts synthesized ECG quality, we alter the denoising steps from 2 to 100 and observe the variation of ECG accuracy (PCC with ground truth). It is noted that during training, the

Table 3. User-independent Performance w/wo calibration guidance (CG).

Experiment	PCC(N)	PCC(D)	RMSE(N)	RMSE(D)
with CG	0.853	0.819	0.182	0.290
without CG	0.794	0.741	0.211	0.329

denoising is configured at fixed 1000 steps, while during evaluation, the count of denoising steps can be altered in the range of 2 to 1000 by interpolation. The PCC variation against denoising step is plotted in Fig. 16(b), where the PCC shows higher improvement as more denoising steps are performed. For disease monitoring, the PCC increases from 0.656 to 0.835 in the preceding 25 steps, while the following increment until 100 steps (0.860) is smaller. However, the computing resources required for inference increase linearly with steps added, thus an appropriate number of denoising steps can be selected based on hardware computing resources. In the setting of *AirECG*, 100 steps are set as default.

**Impact of calibration guidance.** The calibration guidance is the module to control the randomness of diffusion generative model, thus *AirECG* outputs ECG with higher accuracy. To evaluate how the calibration guidance impacts ECG synthesis, we perform an ablation study on the user-independent datasets. Since for user-dependent datasets, the personal ECG features in the test set have been learned during training process (same subject in both train and test). While for the user-independent experiment, the calibration guidance can provide the guidance for the unseen subject in the test set. Table 3 shows the comparison of *AirECG*'s performance with/without calibration guidance. It can be summarized that, (i) The calibration guidance can improve the accuracy of both in trend (PCC) and amplitude (RMSE), showing the effectiveness to control the ECG synthesis process with reference ECG. (ii) Despite the lack of calibration guidance, *AirECG* still outperforms the baselines in Table 1. In particular, the PCC of normal subjects improves from 0.754, 0.755 to 0.794, and PCC of disease subjects improves from 0.643, 0.654 to 0.741. Such improvement implies that the cross-domain diffusion provides better ECG monitoring despite no reference ECG for calibration.

## 7 Discussion and Future Work

### 7.1 *AirECG* on Unseen Data

The ECG synthesis performance on unseen data is a critical issue for the generalization of *AirECG*, since it is a machine learning based system. Despite the mmWave data in the test set having never appeared in the training set during all the experiment settings, we can still divide the unseen data issue into three gradually harder levels. (i) The target disease and user both have been learned from the training set, that is the user-dependent experiment in Sec. 6.1. (ii) The target user is not included in the training data, but the target disease has been learned by *AirECG* using data from previous patients. Such setting has been evaluated as the user-independent experiment in Sec. 6.2. (iii) For the further extreme situations, both target disease and user are not included in the training set and thus are unknown for the model. *AirECG*'s performance under (i) and (ii) conditions has been validated in the previous section, and we will further discuss the unknown diseases issue in the following paragraphs.

In practical scenarios, *AirECG* can be trained on complete disease datasets, thus it can better serve the ECG monitoring tasks of various diseases. In this section, we aim to investigate *AirECG*'s performance on unknown diseases. To simulate such a condition, we delete specific diseases from the training datasets, such as atrial fibrillation (AF) and premature ventricular contraction (PVC). Then AF and PVC data are employed for performance evaluation as model's unknown diseases.

Fig. 17 illustrates *AirECG*'s PCC performance on AF and PVC when the model has seen (Seen) or has not seen (Unseen) the disease during the training process. As we mentioned in Sec. 6.1, PCC of the general training process (Seen) on both diseases shows strong correlation at 0.802 of AF, 0.959 of PVC. However, for the model

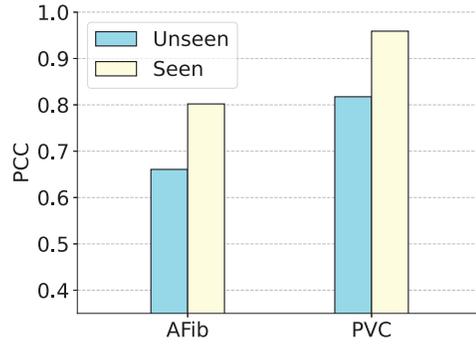


Fig. 17. PCC of a specific disease when it is not included in the training datasets.

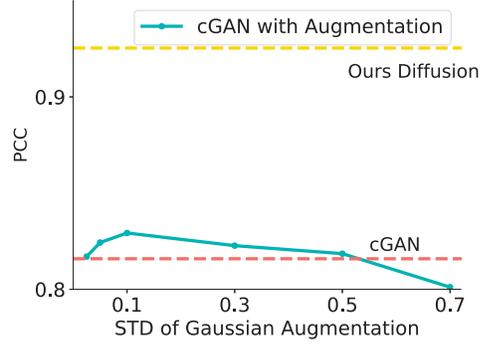


Fig. 18. Performance of conventional cGAN with various Gaussian Augmentation.

unseen specific diseases, the PCC of AF drops to 0.661 and that of PVC drops to 0.818. We can summarize that, (i) For unknown diseases, the consistency between *AirECG* and ECG Patch declines due to the ambiguous mapping between mechanical and electrical activities. (ii) The unseen model performance varies among diseases. For arrhythmia that still has a certain regularity (like PVC), *AirECG* can still maintain a strong correlation with ground truth. But for absolute irregularity of the heart rhythm (like AF), the PCC is unsatisfactory. In conclusion, it is necessary to train *AirECG* on multiple types of diseases, so as to improve the accuracy and generalization ability.

## 7.2 How Diffusion Beats Conventional Model

Whether in our ECG synthesis tasks or in popular image generation tasks, diffusion models show better performance than prior generative models. In this section, we discuss how the denoising diffusion beats conventional models in terms of contactless ECG synthesis. The idea of adding noise and denoising appears to be similar with data augmentation through more samples, while the latter is a common approach to improve the performance of machine learning models. To prove the differences between diffusion and noise augmentation, we employ the conditional GAN (cGAN) in prior RF-ECG [42], and add Gaussian noise into the mmWave data during model training so as to achieve augmentation.

Fig 18 plots the performance variation with different scales of Gaussian noise added to mmWave for augmentation. We set the mean of Gaussian noise to zero, and the standard deviation(STD) ranges from 0.025 to 0.7, since the input data is normalized in  $[-1,1]$ . The red dashed line indicates the original performance of cGAN, showing PCC with ground truth at 0.816, and the yellow dashed line shows diffusion performance (*AirECG*) at 0.925 on the overall test datasets. With the Gaussian noise augmentation at various scales, the PCC of cGAN rises to 0.829 when  $STD=0.1$  and declines to 0.801 when  $STD=0.7$ . It can be summarized that (i) Diffusion model shows better performance than conventional GAN model even if the GAN is equipped with noise augmentation. (ii) Different from randomly adding noise for augmentation, diffusion is a probabilistic model to perform denoising synthesis step-by-step, which means that the Gaussian sampling steps are in serial and directional, rather than isolated steps for data augmentation. (iii) Another weakness of GAN is the problem of mode collapse [17], which means that the balance between generator and discriminator is hard to maintain, and leads to unsatisfactory model generalization. However, diffusion model performs stable synthesis through multiple iteration inference in serial, avoiding the use of a discriminator and thus preventing the problem of mode collapse.

### 7.3 Future Work

**Multi-lead ECG monitoring.** *AirECG* achieves contactless single-lead ECG for disease monitoring, so as to enable arrhythmia detection and management. *AirECG*'s single-lead ECG signal is adequate for heart rhythm disease monitoring (AF, tachycardia and *etc.*), while for ischemic heart disease such as myocardial ischemia, multi-lead ECG is still required for further inspection. The cooperation and comparison between multi-lead ECG signals can boost the ability to detect myocardial ischemia and find out ischemic location. A 12-lead ECG device can monitor chest (precordial) lead, which can locate ventricular issues in horizontal plane. *AirECG* has the potential for multi-lead ECG monitoring since the mmWave radar captures vibration of the overall chest, which covers the positions of multi-lead electrodes. However, the ground truth ECG device in our experiment is a single-lead ECG patch [4], which limits multi-lead ECG synthesis. In the future work, we will try to take advantage of multi-lead ECG devices, especially 12-lead ECG devices in order to obtain the ground truth for cross-domain diffusion training and evaluation.

**Disease severity evaluation and management.** In our experiment, the *AirECG* has been implemented for ECG monitoring in clinical or laboratory setting lasting several hours. For the subsequent research, we plan to extend *AirECG* into longer-term scenarios such as daily monitoring in the bedroom, office and hospital ward. With the help of longer-term daily monitoring, further disease severity evaluation and management application can be investigated. For instance, precise planning of medication usage for paroxysmal AF can be achieved through daily ECG monitoring [28], which benefits in controlling the side effects of medication. Similar application has been validated with contactless breath monitoring. [44] has employed mmWave radar to perform long-term breath monitoring on Parkinson's disease and provided patients with a better medication plan. Additionally, we will improve the noise cancelling ability of *AirECG* to deal with motion artifacts during daily monitoring.

## 8 Related Work

Our work focuses on cardiac sensing topics, and we separate related works into two classes: wearable cardiac sensing and the emerging contactless cardiac sensing.

**Wearable cardiac sensing.** (i) Photoplethysmogram (PPG) is a wearable device-friendly technique for heart monitoring, which monitors blood volume changes at fingertip or on the wrist. [48] propose a tree boosting model to analyze heart rate variability (HRV) from PPG data, while [36, 46] utilize light weight CNN with PPG data from smartwatch to extract HRV features for atrial fibrillation (AF) screening. However, PPG can not provide the atrial and ventricular activities in the electrical domain, which limits the cardiac monitoring accuracy and further medical applications. (ii) ECG has been widely used for clinical diagnosis and monitoring of CVDs, such as arrhythmia, heart structure changes, *etc.* [5]. In recent studies, portable and wearable ECG devices have been developed to realize long-term monitoring and mobile health applications. Miniaturized patch ECG devices like Zio monitor [45] and CarePulse [4] can provide single-lead ECG through attaching to the anterior chest, and these devices have demonstrated their ability to detect and classify arrhythmias such as atrial fibrillation (AF) and premature beats [22] with residual CNN. However, they are still conventional ECG devices requiring electrode patches, which may lead to allergic skin problems and medical waste. Smartwatches are another type of wearable devices for ECG monitoring, achieving AF [24] and ischemia [9] detection in recent works. However, they do not support long-term monitoring due to the requirement of actively touching the electrode.

**Contactless cardiac sensing.** Wireless signals, including WiFi, ultra wide band (UWB) and mmWave have been investigated for contactless cardiac health monitoring in recent years, mainly through capturing the thorax wall vibration to extract mechanical activities of hearts. (i) Early works deploy WiFi [40, 41], UWB [13] and mmWave radar [39, 49] to monitor heart rate and its variability. Based on heart rate monitoring, a variety of contactless cardiac sensing applications can be realized unobtrusively, such as emotion recognition [49] and stress level estimation[21]. [20] further attempts to capture the fine-grained atrial and ventricular vibration

called seismocardiography (SCG) using mmWave radar, which is realized through multichannel beamforming and 1D-CNN. In our early work [47], we propose an IQ-domain arrhythmia detector for mmWave radar signal, which is an end-to-end arrhythmia monitoring system with raw IQ signal input. (ii) Since the electrical activities of heart are the gold standard in medical usage, recent studies have focused on contactless ECG monitoring [11, 42, 43], which take advantages of generative models to transform wireless signals into ECG data. For instance, CardiacWave [43] designs an attention and LSTM based CaSE-ECG Solver to reconstruct ECG-like signal via Cardiac-mmWave scattering effect. RF-ECG [42] proposes a cGAN model to generate ECG signal based on the condition from UWB cardiac sensing.

Different from prior works, *AirECG* promotes the realization of contactless ECG monitoring for real patients with cardiac diseases. More specifically, we try to realize ECG synthesis even for irregular and abnormal heartbeats, which can be achieved through multiple iterations inference to improve the deep learning model. In this way, *AirECG* can implement accurate and robust cardiac disease monitoring in practical scenarios.

## 9 Conclusion

In this paper, we propose *AirECG*, a mmWave-based contactless ECG system that can be used for disease monitoring in daily scenarios. We custom-design a cross-domain diffusion model that can transfer mechanical domain chest vibration into electrical domain ECG, which is achieved through multiple iterations of denoising inference, so as to accurately synthesize ECG wave even during abnormal heartbeats. In addition, we enhance the robustness of *AirECG* by ECG calibration guidance, which helps to address the variability in ECG features caused by probabilistic diffusion model and individual differences. Our experiment illustrates that *AirECG* achieves compelling results for ECG monitoring, which implies a significant improvement over previous contactless ECG works. We believe that *AirECG* promotes a meaningful step towards contactless ECG, which is promising to realize disease screening, severity evaluation and management in both healthy individuals and patients.

## Acknowledgments

We appreciate the insightful feedback from the editors and reviewers who helped improve this work. This work was supported in part by Beijing Natural Science Foundation(L223002), the Innovation Research Group Project of NSFC under Grant 61921003, the Youth Top Talent Support Program.

## References

- [1] 2023. ECG Analysis Platform of Carepulse. <https://platform.carepulse.cn:9443/login>. Accessed July 9, 2023.
- [2] 2024. Atrial Fibrillation. <https://www.mayoclinic.org/diseases-conditions/atrial-fibrillation/symptoms-causes/syc-20350624>. Accessed Jan 15, 2024.
- [3] 2024. Cardiovascular diseases. <https://www.who.int/health-topics/cardiovascular-diseases/>. Accessed Jan 15, 2024.
- [4] 2024. Carepulse Patch. <https://index.carepulse.cn/home/index.html>. Accessed Jan 15, 2024.
- [5] 2024. Electrocardiogram (ECG or EKG). <https://www.mayoclinic.org/tests-procedures/ekg/about/pac-20384983/>. Accessed Jan 15, 2024.
- [6] 2024. Holter Monitor. [https://en.wikipedia.org/wiki/Holter\\_monitor](https://en.wikipedia.org/wiki/Holter_monitor). Accessed Jan 15, 2024.
- [7] 2024. Texas Instruments DCA1000EVM. <https://www.ti.com/tool/DCA1000EVM>. Accessed Jan 15, 2024.
- [8] 2024. Texas Instruments IWR1443BOOST. <https://www.ti.com/tool/IWR1443BOOST>. Accessed Jan 15, 2024.
- [9] Cesar O Avila. 2019. Novel use of Apple Watch 4 to obtain 3-lead electrocardiogram and detect cardiac ischemia. *The Permanente Journal* 23 (2019).
- [10] Hugh Calkins, Matthew R Reynolds, Peter Spector, Manu Sondhi, Yingxin Xu, Amber Martin, Catherine J Williams, and Isabella Sledge. 2009. Treatment of atrial fibrillation with antiarrhythmic drugs or radiofrequency ablation: two systematic literature reviews and meta-analyses. *Circulation: Arrhythmia and Electrophysiology* 2, 4 (2009), 349–361.
- [11] Jinbo Chen, Dongheng Zhang, Zhi Wu, Fang Zhou, Qibin Sun, and Yan Chen. 2022. Contactless electrocardiogram monitoring with millimeter wave radar. *IEEE Transactions on Mobile Computing* (2022).
- [12] Nanxin Chen, Yu Zhang, Heiga Zen, Ron J Weiss, Mohammad Norouzi, and William Chan. 2020. Wavegrad: Estimating gradients for waveform generation. *arXiv preprint arXiv:2009.00713* (2020).

- [13] Zhe Chen, Tianyue Zheng, Chao Cai, and Jun Luo. 2021. MoVi-Fi: Motion-robust vital signs waveform recovery via deep interpreted RF sensing. In *Proceedings of the 27th annual international conference on mobile computing and networking*. 392–405.
- [14] Guoxuan Chi, Zheng Yang, Chenshu Wu, Jingao Xu, Yuchong Gao, Yunhao Liu, and Tony Xiao Han. 2024. RF-Diffusion: Radio Signal Generation via Time-Frequency Diffusion. In *Proceedings of the 30th Annual International Conference on Mobile Computing and Networking*. 77–92.
- [15] Guoxuan Chi, Guidong Zhang, Xuan Ding, Qiang Ma, Zheng Yang, Zhenguo Du, Houfei Xiao, and Zhuang Liu. 2024. XFall: Domain Adaptive Wi-Fi-based Fall Detection with Cross-Modal Supervision. *IEEE Journal on Selected Areas in Communications* (2024).
- [16] Anne De Groot, Muriel Wantier, Guy Chéron, Marc Estenne, and Manuel Paiva. 1997. Chest wall motion during tidal breathing. *Journal of Applied Physiology* 83, 5 (1997), 1531–1537.
- [17] Prafulla Dhariwal and Alexander Nichol. 2021. Diffusion models beat gans on image synthesis. *Advances in neural information processing systems* 34 (2021), 8780–8794.
- [18] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, et al. 2020. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929* (2020).
- [19] Ary L Goldberger, Luis AN Amaral, Leon Glass, Jeffrey M Hausdorff, Plamen Ch Ivanov, Roger G Mark, Joseph E Mietus, George B Moody, Chung-Kang Peng, and H Eugene Stanley. 2000. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *circulation* 101, 23 (2000), e215–e220.
- [20] Unsoo Ha, Salah Assana, and Fadel Adib. 2020. Contactless seismocardiography via deep learning radars. In *Proceedings of the 26th annual international conference on mobile computing and networking*. 1–14.
- [21] Unsoo Ha, Sohrab Madani, and Fadel Adib. 2021. WiStress: Contactless stress monitoring using wireless signals. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 3 (2021), 1–37.
- [22] Awni Y Hannun, Pranav Rajpurkar, Masoumeh Haghpanahi, Geoffrey H Tison, Codie Bourn, Mintu P Turakhia, and Andrew Y Ng. 2019. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nature medicine* 25, 1 (2019), 65–69.
- [23] Jonathan Ho, Ajay Jain, and Pieter Abbeel. 2020. Denoising diffusion probabilistic models. *Advances in neural information processing systems* 33 (2020), 6840–6851.
- [24] Nino Isakadze and Seth S Martin. 2020. How useful is the smartwatch ECG? *Trends in cardiovascular medicine* 30, 7 (2020), 442–448.
- [25] Wen-Yen Lin, Wen-Cheng Chou, Po-Cheng Chang, Chung-Chuan Chou, Ming-Shien Wen, Ming-Yun Ho, Wen-Chen Lee, Ming-Jer Hsieh, Chung-Chih Lin, Tsai-Hsuan Tsai, et al. 2016. Identification of location specific feature points in a cardiac cycle using a novel seismocardiogram spectrum system. *IEEE journal of biomedical and health informatics* 22, 2 (2016), 442–449.
- [26] José A López-López, Jonathan AC Sterne, Howard HZ Thom, Julian PT Higgins, Aroon D Hingorani, George N Okoli, Philippa A Davies, Pritesh N Bodalia, Peter A Bryden, Nicky J Welton, et al. 2017. Oral anticoagulants for prevention of stroke in atrial fibrillation: systematic review, network meta-analysis, and cost effectiveness analysis. *bmj* 359 (2017).
- [27] Dominique Makowski, Tam Pham, Zen J Lau, Jan C Brammer, François Lespinasse, Hung Pham, Christopher Schölzel, and SH Annabel Chen. 2021. NeuroKit2: A Python toolbox for neurophysiological signal processing. *Behavior research methods* (2021), 1–8.
- [28] David T Martin, Malcolm M Bersohn, Albert L Waldo, Mark S Wathen, Wassim K Choucair, Gregory YH Lip, John Ip, Richard Holcomb, Joseph G Akar, and Jonathan L Halperin. 2015. Randomized trial of atrial arrhythmia monitoring to guide anticoagulation in patients with implanted defibrillator and cardiac resynchronization devices. *European heart journal* 36, 26 (2015), 1660–1668.
- [29] George B Moody and Roger G Mark. 2001. The impact of the MIT-BIH arrhythmia database. *IEEE engineering in medicine and biology magazine* 20, 3 (2001), 45–50.
- [30] Jingyi Ning, Lei Xie, Chuyu Wang, Yanling Bu, Fengyuan Xu, Da-Wei Zhou, Sanglu Lu, and Baoliu Ye. 2021. RF-badge: vital sign-based authentication via RFID tag array on badges. *IEEE Transactions on Mobile Computing* (2021).
- [31] William Peebles and Saining Xie. 2023. Scalable diffusion models with transformers. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 4195–4205.
- [32] Emily R Pfeiffer, Jared R Tangney, Jeffrey H Omens, and Andrew D McCulloch. 2014. Biomechanics of cardiac electromechanical coupling and mechanoelectric feedback. *Journal of biomechanical engineering* 136, 2 (2014), 021007.
- [33] SP Preejith, R Dhinesh, Jayaraj Joseph, and Mohanasankar Sivaprakasam. 2016. Wearable ECG platform for continuous cardiac monitoring. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 623–626.
- [34] Aditya Ramesh, Prafulla Dhariwal, Alex Nichol, Casey Chu, and Mark Chen. 2022. Hierarchical text-conditional image generation with clip latents. *arXiv preprint arXiv:2204.06125* 1, 2 (2022), 3.
- [35] Gregory A Roth, Catherine O Johnson, Kalkidan Hassen Abate, Foad Abd-Allah, Muktar Ahmed, Khurshid Alam, Tahiya Alam, Nelson Alvis-Guzman, Hossein Ansari, Johan Ärnlöv, et al. 2018. The burden of cardiovascular diseases among US states, 1990–2016. *JAMA cardiology* 3, 5 (2018), 375–389.
- [36] Yichen Shen, Maxime Voisin, Alireza Aliamiri, Anand Avati, Awni Hannun, and Andrew Ng. 2019. Ambulatory atrial fibrillation monitoring using wearable photoplethysmography with deep learning. In *Proceedings of the 25th ACM SIGKDD International Conference*

- on *Knowledge Discovery & Data Mining*. 1909–1916.
- [37] Yang Song and Stefano Ermon. 2019. Generative modeling by estimating gradients of the data distribution. *Advances in neural information processing systems* 32 (2019).
- [38] Roger Stevens. 2006. *Gray’s Anatomy for Students*.
- [39] Fengyu Wang, Xiaolu Zeng, Chenshu Wu, Beibei Wang, and KJ Ray Liu. 2021. mmHRV: Contactless heart rate variability monitoring using millimeter-wave radio. *IEEE Internet of Things Journal* 8, 22 (2021), 16623–16636.
- [40] Xuyu Wang, Chao Yang, and Shiwen Mao. 2017. PhaseBeat: Exploiting CSI phase data for vital sign monitoring with commodity WiFi devices. In *2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS)*. IEEE, 1230–1239.
- [41] Xuyu Wang, Chao Yang, and Shiwen Mao. 2020. On CSI-based vital sign monitoring using commodity WiFi. *ACM Transactions on Computing for Healthcare* 1, 3 (2020), 1–27.
- [42] Zhi Wang, Beihong Jin, Siheng Li, Fusang Zhang, and Wenbo Zhang. 2023. ECG-grained Cardiac Monitoring Using UWB Signals. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 4 (2023), 1–25.
- [43] Chenhan Xu, Huining Li, Zhengxiong Li, Hanbin Zhang, Aditya Singh Rathore, Xingyu Chen, Kun Wang, Ming-chun Huang, and Wenyao Xu. 2021. Cardiacwave: A mmwave-based scheme of non-contact and high-definition heart activity computing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 3 (2021), 1–26.
- [44] Yuzhe Yang, Yuan Yuan, Guo Zhang, Hao Wang, Ying-Cong Chen, Yingcheng Liu, Christopher G Tarolli, Daniel Crepeau, Jan Bukartyk, Mithri R Junna, et al. 2022. Artificial intelligence-enabled detection and assessment of Parkinson’s disease using nocturnal breathing signals. *Nature medicine* 28, 10 (2022), 2207–2215.
- [45] Mihran Yenikomshian, John Jarvis, Cody Patton, Christopher Yee, Richard Mortimer, Howard Birnbaum, and Mark Topash. 2019. Cardiac arrhythmia detection outcomes among patients monitored with the Zio patch system: a systematic literature review. *Current Medical Research and Opinion* 35, 10 (2019), 1659–1670.
- [46] Hanbin Zhang, Li Zhu, Viswam Nathan, Jilong Kuang, Jacob Kim, Jun Alex Gao, and Jeffrey Olgin. 2021. Towards early detection and burden estimation of atrial fibrillation in an ambulatory free-living environment. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 2 (2021), 1–19.
- [47] Langcheng Zhao, Rui Lyu, Qi Lin, Anfu Zhou, Huanhuan Zhang, Huadong Ma, Jingjia Wang, Chunli Shao, and Yida Tang. 2024. mmArrhythmia: Contactless Arrhythmia Detection via mmWave Sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8, 1 (2024), 1–25.
- [48] Langcheng Zhao, Fenglin Zhang, Huanhuan Zhang, Yumeng Liang, Anfu Zhou, and Huadong Ma. 2022. Robust respiratory rate monitoring using smartwatch photoplethysmography. *IEEE Internet of Things Journal* 10, 6 (2022), 4830–4844.
- [49] Mingmin Zhao, Fadel Adib, and Dina Katabi. 2016. Emotion recognition using wireless signals. In *Proceedings of the 22nd annual international conference on mobile computing and networking*. 95–108.