Radar-APLANC: Unsupervised Radar-based Heartbeat Sensing via Augmented Pseudo-Label and Noise Contrast

Ying Wang[1], **Zhaodong Sun**[1*], **Xu Cheng**[1], **Zuxian He**[1], **Xiaobai Li**[2]
[1]Nanjing University of Information Science & Technology, Nanjing, China
[2]Zhejiang University, Hangzhou, China
{ying.wang, zhaodong.sun, xcheng, hezuxian}@nuist.edu.cn, xiaobai.li@zju.edu.cn

Abstract

Frequency Modulated Continuous Wave (FMCW) radars can measure subtle chest wall oscillations to enable non-contact heartbeat sensing. However, traditional radar-based heartbeat sensing methods face performance degradation due to noise. Learning-based radar methods achieve better noise robustness but require costly labeled signals for supervised training. To overcome these limitations, we propose the first unsupervised framework for radar-based heartbeat sensing via **A**ugmented **P**seudo-**La**bel and **N**oise **C**ontrast (Radar-APLANC). We propose to use both the heartbeat range and noise range within the radar range matrix to construct the positive and negative samples, respectively, for improved noise robustness. Our Noise-Contrastive Triplet (NCT) loss only utilizes positive samples, negative samples, and pseudo-label signals generated by the traditional radar method, thereby avoiding dependence on expensive ground-truth physiological signals. We further design a pseudo-label augmentation approach featuring adaptive noise-aware label selection to improve pseudo-label signal quality. Extensive experiments on the Equipleth dataset and our collected radar dataset demonstrate that our unsupervised method achieves performance comparable to state-of-the-art supervised methods.

Code —
https://github.com/RadarHRSensing/Radar-APLANC

Introduction

Radar-based heartbeat sensing has emerged as a pivotal technology for non-contact physiological assessment, offering distinct advantages in privacy preservation, environmental robustness, and continuous physiological monitoring. The basic principle for radar-based heartbeat sensing is detecting sub-millimeter chest wall displacements (typically 0.1–0.5 mm) induced by cardiac subtle motion (Droitcour 2006). Frequency Modulated Continuous Wave (FMCW) radars are widely used to capture heartbeat signals by measuring the relative phase changes of the received chirp signals. Traditional approaches directly extract and unwrap the phases of received chirp signals (Alizadeh et al. 2019; Tu, Hwang, and Lin 2016; Mercuri et al. 2019). However, these traditional methods suffer significant performance degradation under motion artifacts, multipath interference, and low signal-to-noise conditions due to inherent phase wrapping ambiguities and noise sensitivity as shown in Fig. 1(a).

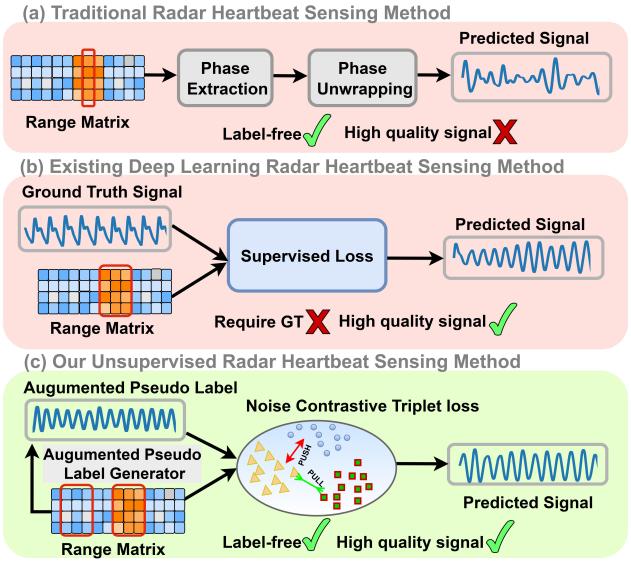

Figure 1: Comparison of radar-based heartbeat sensing methods. (a) Traditional methods simply rely on phase extraction and unwrapping processing, resulting in low signal quality; (b) Existing supervised methods require ground truth signals; (c) Our proposed unsupervised approach generates high-quality predictions without requiring ground truth signals.

Recent supervised deep learning methods (Vilesov et al. 2022; Hu et al. 2024; Wu et al. 2025) have demonstrated improved robustness by learning complex spatiotemporal patterns directly from radar data. However, these approaches require large-scale datasets with high-quality physiological annotations (e.g., synchronized PPG signals), which are costly to acquire, as shown in Fig. 1(b). This dependency creates the fundamental limitation for supervised radar methods: scalability bottlenecks of training data due to the expensive physiological annotations.

While unsupervised learning paradigms have shown

*Corresponding author.

promise in video-based physiological monitoring (Gideon and Stent 2021; Sun and Li 2022; Speth et al. 2023; Li and Yin 2023; Yue, Shi, and Ding 2023), directly adapting these to radar data faces inherent incompatibility and obstacles. Heartbeat signals in radar data usually exhibit lower signal-to-noise ratios compared to the video modality (Vilesov et al. 2022); heartbeat signals manifest differently between radar versus video domains, i.e., chest motions for radar and facial color changes for videos; conventional positive and negative pair construction strategies (e.g., spatiotemporal similarity cross-sample dissimilarity) (Sun and Li 2022) fail under strong noise interference in radar data. These challenges indicate that a specialized unsupervised framework should be designed for radar-based heartbeat sensing.

To address these challenges, we propose Radar-APLANC: a two-stage unsupervised radar-based heartbeat sensing framework built on augmented pseudo-label and noise contrast (APLANC) as shown in Fig. 1(c). Our approach introduces two key innovations: **1)** a Noise-Contrastive Triplet (NCT) loss is proposed by contrasting heartbeat signals with pseudo-label signals and noise signals from a radar range matrix; **2)** an augmented pseudo-label generator that refines pseudo-labels through quality assessment and adaptive noise-ware label selection. This two-stage framework enables effective unsupervised learning without physiological labels while maintaining noise robustness.

Comprehensive evaluations on the Equipleth dataset (Vilesov et al. 2022) and our collected radar heartbeat (RHB) dataset demonstrate that Radar-APLANC achieves close performance to existing state-of-the-art supervised methods. The results validate our unsupervised framework's capability to overcome fundamental limitations in traditional or supervised radar methods for heartbeat sensing.

The main contributions can be summarized as follows:

1. We pioneer the first unsupervised framework for radar-based heartbeat sensing, eliminating dependency on physiological labels and achieving comparable performance with supervised methods.
2. We propose the NCT loss – the first attempt to exploit noise artifacts in radar range matrices to improve the noise robustness for radar-based heartbeat sensing.
3. We develop a two-stage training strategy with an augmented pseudo-label generator featuring adaptive noise-aware label selection from radar range matrices.
4. We collected a new radar-based heartbeat sensing benchmark dataset (RHB) from 80 subjects, which will be open-sourced for community research.

Related Work

Radar-based Heartbeat Sensing

Noncontact Radar-based heartbeat measurement has gained attention in recent works. (Alizadeh et al. 2019) combines phase unwrapping with range-FFT to detect heartbeat and respiration signals using FMCW radar in real-world settings. (Mercuri et al. 2019) introduces a radar system capable of simultaneously monitoring vital signs and tracking the spatial positions of multiple individuals without physical contact. (Tu, Hwang, and Lin 2016) presents a method that enables accurate respiration rate estimation under one-dimensional body motion using a single continuous-wave Doppler radar with motion compensation.

Recent efforts (Hu et al. 2024; Vilesov et al. 2022; Wu et al. 2025; Khan, Rigazio, and Shahzad 2022) apply supervised deep learning to estimate heartbeat signals from radar data, achieving strong results by modeling temporal patterns. However, these methods require large annotated radar datasets, limiting training and data scalability.

To address this, recent studies (Song et al. 2022; Zhang et al. 2025; Song et al. 2024) explore self-supervised learning for radar-based sensing tasks, including human gesture recognition, 3D pose estimation, silhouette generation, and electrocardiography (ECG) signal reconstruction. While these studies reduce dependence on labeled data, they still require some annotations for fine-tuning and cannot achieve unsupervised learning.

Unsupervised Learning with Pseudo-labels

Pseudo-labeling has been widely used in unsupervised learning, especially for classification. Early methods (Saito, Ushiku, and Harada 2017; Chen et al. 2019) improve label quality via multi-classifier refinement or progressive sample selection. Others leverage generative models (Morerio et al. 2020; Wang et al. 2022) or contrastive strategies (Sun et al. 2023; Diamant et al. 2024; Litrico, Del Bue, and Morerio 2023) to address noisy labels in domain adaptation and source-free settings. Curriculum learning and debiasing approaches (Choi et al. 2019; Lai et al. 2023) enhance pseudo-labeling by adjusting label difficulty and correcting imbalance during training. In cross-modal tasks like visible-infrared person re-ID, clustering-based pseudo-labeling (Shi et al. 2024) and soft assignment (Lin et al. 2020; Seo, Lee, and Han 2022) further boost performance. However, these techniques are tailored for classification and are less applicable to regression tasks like heartbeat signal measurement. (Li and Yin 2023) utilizes traditional video-based methods to estimate heartbeat signals as pseudo-labels to achieve unsupervised training. Inspired by this work, our method initially generates pseudo-labels using traditional radar-based heartbeat estimation (Alizadeh et al. 2019) as weak label supervision, but the noisy radar pseudo-labels are suboptimal. Therefore, we further incorporate noise information and design an augmented pseudo-label generator to adaptively select the radar pseudo-labels for improved performance.

Method

In this section, we will first introduce the preliminaries of radar-based heartbeat sensing as the basis for understanding our proposed method. Subsequently, we introduce our two-stage unsupervised method (Radar-APLANC) for radar-based heartbeat sensing. The overviews of the two stages are shown in Fig. 2 and Fig. 3, respectively.

Preliminaries

Range Matrix. A range matrix is obtained from FMCW radar raw data to facilitate the following analysis and pro-

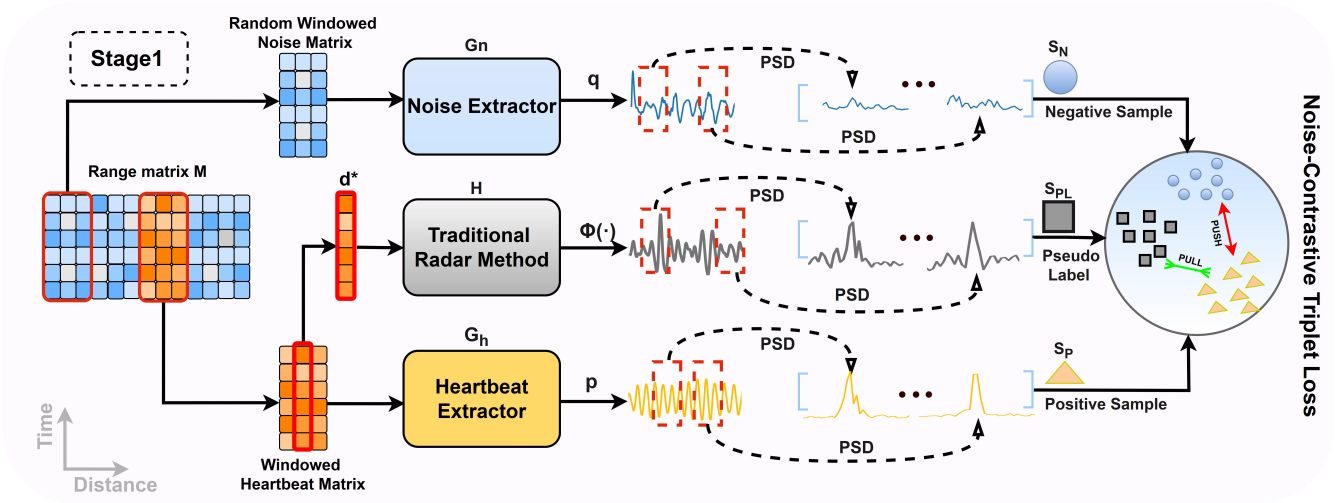

Figure 2: The framework of Stage One. The heartbeat matrix, pseudo-label, and random noise matrix undergo random temporal sampling and power spectrum densities (PSD) transform before being fed into the NCT loss. Within this framework, the PSD of the heart matrix is attracted to that of the pseudo-label while being repelled from the noise PSD.

cessing. The procedures to obtain a range matrix are as follows. An FMCW radar in each chirp loop transmits a chirp signal $s(t)$ and receives the reflected chirp signal $u(t)$. Both $s(t)$ and $u(t)$ are linear frequency modulation signals, also called chirp signals. The received signal $u(t)$ is modulated with the in-phase and quadrature (IQ) transmitting signals $s_I(t)$ and $s_Q(t)$ to get the complex intermediate frequency (IF) signal $m(t) \in \mathbb{R}^D$ as shown below:

$$\begin{aligned} m(t) &\propto \text{LPF}[s_I(t) \cdot u(t)] + j\text{LPF}[s_Q(t) \cdot u(t)] \\ &\propto \exp(j(2\pi f t + \varphi)), f = 2kd/c, \varphi = 4\pi d/\lambda \end{aligned} \quad (1)$$

where LPF is the low-pass filter, k is the frequency slope of FMCW, d is the distance, c is the light speed, λ is the wavelength of the FMCW starting frequency. The frequency f of the IF signal $m(t)$ is the frequency difference between the transmitted signal $s(t)$ and the received signal $u(t)$. The frequency f is also proportional to the signal round-trip time and the distance d between the radar and the object. The phase φ is also proportional to the distance but is bounded between $-\pi$ and π. A radar can sequentially transmit N chirps $[s_1(t), s_2(t), ..., s_N(t)]$ and receive the corresponding N reflected chirps $[u_1(t), u_2(t), ..., u_N(t)]$. In the meanwhile, the radar can obtain N IF signals $[m_1(t), m_2(t), ..., m_N(t)]$. Since the frequency of each IF signal $m_n(t) \in \mathbb{R}^D$ is related to the distance, the fast Fourier transform (FFT) is performed on each IF signal $m_n(t)$ to get the corresponding range profile $M_n[f]$. Finally, all range profiles are concatenated to obtain the range matrix:

$$M = [M_1[f], M_2[f], ..., M_N[f]] \in \mathbb{R}^{N \times D}, \quad (2)$$

where N is the number of chirps, and D is the number of range bins.

Basic Radar-based Heartbeat Sensing. The frequency f cannot be directly used to extract the submillimeter chest motions caused by heartbeats since the range resolution is on the order of centimeters. Instead, the highly sensitive phase φ should be used to measure such submillimeter displacement. The basic heartbeat sensing (Alizadeh et al. 2019) consists of four steps as shown in Fig. 1(a): (1) Select the range bin d^* with the maximum power occupancy along the range axis in the range matrix. This range bin corresponds to the person's position. (2) Calculate the phase angle at the selected range bin for each range profile to get the phase signal $\phi(\cdot) = [\varphi_1, ..., \varphi_N] = [\text{angle}(M_1[d^*]), ..., \text{angle}(M_N[d^*])] \in \mathbb{R}^N$. (3) Since the phases are wrapped between $[-\pi, \pi]$, a standard phase unwrapping algorithm should be used to get the unwrapped phase signal. (4) Filter the unwrapped phase signal with 0.8 Hz-3.0 Hz bandpass to get the heartbeat signal $\Phi(\cdot) \in \mathbb{R}^N$ and find the highest peak corresponding to the heart rate in the frequency domain.

Stage One: Unsupervised Noise-Contrastive Pretraining

Fig. 2 illustrates the framework of stage one, which leverages noise information and heartbeat information to construct pseudo-labels, positive samples, and negative samples for NCT loss. After stage one, our method can primarily extract coarse heartbeat signals.

Pseudo-labels. We use the traditional radar heartbeat sensing steps described above in the preliminaries section to get pseudo-labels. As shown in Fig. 2 for Stage One, we use the traditional radar method (Alizadeh et al. 2019) to extract heartbeat signals from the range bin with the maximum power occupancy. We further perform random temporal sampling and power spectrum densities (PSD) transform shown in Fig. 2 to generate multiple pseudo-labels for subsequent NCT Loss. The pseudo-label set S_{PL} can be described

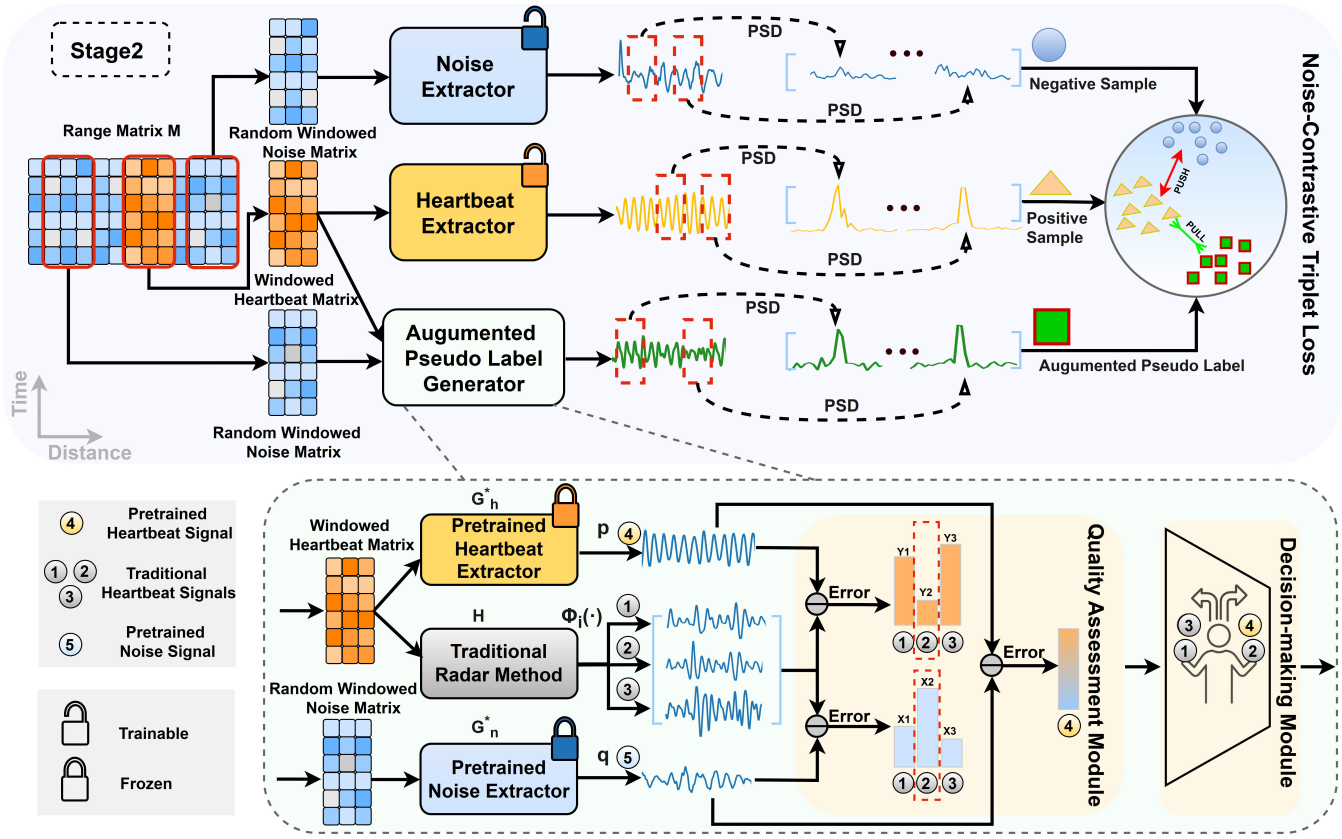

Figure 3: The framework of Stage Two with Augmented Pseudo-label Generator. It consists of Quality Measurement Module and Decision-making Module.

as:

$$S_{PL} = \{\mathcal{P}(\Phi(n_1 \pm \Delta n)), ..., \mathcal{P}(\Phi(n_K \pm \Delta n))\}, \quad (3)$$

where $\mathcal{P}$ means PSD transform, $\Phi(\cdot) \in \mathbb{R}^N$ stands for the heartbeat signal from the traditional method, and $n_k \pm \Delta n$ represents the random time interval around n_k.

Positive Pairs. Following the previous work (Vilesov et al. 2022), we take a window of the range matrix $M \in \mathbb{R}^{N \times D}$ around the central range bin d^* to get the windowed heartbeat matrix $M(\cdot, d^* \pm \Delta d) \in \mathbb{R}^{N \times (2\Delta d+1)}$ as the input to the heartbeat extractor. The extractor can output the predicted heartbeat signal $p(\cdot) \in \mathbb{R}^N$. Similar to pseudo-labels, we perform random temporal sampling and PSD transform on $p(\cdot)$ to get the positive sample set:

$$S_P = \{\mathcal{P}(p(n_1 \pm \Delta n)), ..., \mathcal{P}(p(n_K \pm \Delta n))\} \quad (4)$$

Since pseudo-labels contain heartbeat information, pseudo-labels and the predicted heartbeat signals should be similar, and the samples in the pseudo-label set S_{PL} and the heartbeat signal set S_P should be pulled together as positive pairs. Therefore, the positive term of the NCT loss can be represented as follows:

$$\mathcal{L}_P = \frac{1}{K^2} \sum_{i=1}^{K} \sum_{j=1}^{K} ||S_{PL}[i] - S_P[j]||^2 \quad (5)$$

Negative Pairs. We form negative samples using random windowed noise matrices extracted from the range matrix. Since heartbeat information is primarily concentrated in the heartbeat matrix, other range bins mainly contain background noise. Therefore, we randomly select the range bin d' except the central range bin d^*. Similar to the windowed heartbeat matrix, the windowed noise matrix is $M(\cdot, d' \pm \Delta d)$ as the input to the noise extractor. The noise extractor estimates the noise signal $q(\cdot) \in \mathbb{R}^N$. After temporal sampling and PSD transform, we get the negative sample set:

$$S_N = \{\mathcal{P}(q(n_1 \pm \Delta n)), ..., \mathcal{P}(q(n_K \pm \Delta n))\}. \quad (6)$$

Since heartbeat signals should not contain such background noises, the negative term of the NCT loss can be formulated as follows:

$$\mathcal{L}_N = -\frac{1}{K^2} \sum_{i=1}^{K} \sum_{j=1}^{K} ||S_P[i] - S_N[j]||^2 \quad (7)$$

The overall NCT loss is the sum of positive and negative loss terms: $\mathcal{L}_{NCT} = \mathcal{L}_P + \mathcal{L}_N$. The detailed training procedures of stage one are presented in Algorithm 1.

Stage Two: Unsupervised Learning with Augmented Pseudo-label Generation

Due to the significant noise inherent in radar data, the performance in stage one, using the traditional radar method to

Algorithm 1: Radar-APLANC Training Algorithm

Input: Range matrix dataset $\{M_1, ...M_Q\}$; Heartbeat extractor G_h; Noise extractor G_n; Pretrained heartbeat extractor (only for stage two) G_h^*; Pretrained noise extractor (only for stage two) G_n^*; Traditional radar method H; Training stage STAGE.

for e = 1:NumIteration **do**
  Load a range matrix M_i.
  For the positive sample
  Find the central range bin d^* in M_i.
  $p = G_h(M_i(\cdot, d^* \pm \Delta d)) \leftarrow$ predicted heartbeat signal
  For the negative sample
  Randomly select a range bin d' except d^* in M_i.
  $q = G_n(M_i(\cdot, d' \pm \Delta d)) \leftarrow$ noise signal
  For the pseudo-label
  if STAGE==1 **then**
    $\Phi = H(M_i(\cdot, d^*)) \leftarrow$ traditional method
  else if STAGE==2 **then**
    Φ = AUGPSEUDOGEN(M_i, d^*, G_h^*, G_n^*)
  end if
  Calculate NCT loss and update G_h and G_n.
end for
Augmented Pseudo-label Generator
function AUGPSEUDOGEN(M, d^*, G_h^*, G_n^*)
  $\Phi_1, ..., \Phi_{2\Delta d+1} = H(M(\cdot, d^* \pm \Delta d)) \leftarrow$ traditional heartbeat
  $p = G_h^*(M(\cdot, d^* \pm \Delta d)) \leftarrow$ pretrained heartbeat
  Randomly select a range bin d' except d^* in M.
  $q = G_n^*(M(\cdot, d' \pm \Delta d)) \leftarrow$ pretrained noise
  Noise distance X_i and heartbeat distance Y_i
  $X_i, Y_i = D(\Phi_i, q), D(\Phi_i, p), i \in [1, 2\Delta d + 1]$
  Decision-making Module
  if $\arg\max_i X_i == \arg\min_i Y_i$ **then**
    Return $\Phi_{\arg\min_i Y_i}$
  else
    if $X_{\arg\min_i Y_i} > D(p, q)$ **then**
      Return $\Phi_{\arg\min_i Y_i}$
    else
      Return p
    end if
  end if
end function

generate pseudo-labels, remains limited. Therefore, we introduce an augmented pseudo-label generator to further refine and optimize the selection of pseudo-labels as illustrated in Figure 3.

Candidate Signal Extraction. We extract a pretrained heartbeat signal p from the windowed heartbeat matrix $M(\cdot, d^* \pm \Delta d)$ and a pretrained noise signal q from a random windowed noise matrix $M(\cdot, d' \pm \Delta d)$ using the stage one pretrained extractors. Simultaneously, we extract $2\Delta d + 1$ traditional heartbeat signals $\{\Phi_1, ..., \Phi_{2\Delta d+1}\}$ from the windowed heartbeat matrix $M(\cdot, d^* \pm \Delta d)$ using the traditional radar method. The final pseudo-label will be selected from the traditional heartbeat signals $\{\Phi_1, ..., \Phi_{2\Delta d+1}\}$ and the pretrained heartbeat signal p, while the noise signal q will be used to assess the heartbeat signal qualities.

Quality Assessment Module. We compute two types of distances for candidate heartbeat signals $\{\Phi_1, ..., \Phi_{2\Delta d+1}\}$: (1) Noise distance X_i between candidate heartbeat signals $\{\Phi_1, ..., \Phi_{2\Delta d+1}\}$ and the pretrained noise signal q:

$$X_i = D(\Phi_i, q), i \in [1, 2\Delta d + 1]. \tag{8}$$

Longer noise distances indicate better signal quality. (2) Heartbeat distance Y_i between the traditional heartbeat signals $\{\Phi_1, ..., \Phi_{2\Delta+1}\}$ and the pretrained heartbeat signal p:

$$Y_i = D(\Phi_i, p), i \in [1, 2\Delta d + 1]. \tag{9}$$

which measures how similar the traditional heartbeat signal Φ_i is to the pretrained heartbeat signal p. Since we have an assumption that the pretrained heartbeat signal p has good signal quality, shorter distances indicate better signal quality. To measure the distance between two signals, we use the mean absolute error between the two signals' heart rates.

Decision-making Module. The best heartbeat signal among $\{\Phi_1, ..., \Phi_{2\Delta+1}\}$ should satisfy the condition that its noise distance X_i should be maximum while its heartbeat distance Y_i should be minumum. Therefore, the following equation should hold: $\arg\max_i X_i = \arg\min_i Y_i$. When this ideal condition is not met, i.e., $\arg\max_i X_i \neq \arg\min_i Y_i$, the module evaluates whether the traditional heartbeat signal $\Phi_{\arg\min Y_i}$ with the minimum heartbeat distance is sufficiently distant from the noise q. Specifically, it checks if $\Phi_{\arg\min Y_i}$'s noise distance $X_{\arg\min Y_i}$ is greater than the pretrained heartbeat signal's noise distance $D(p, q)$. If yes, the signal quality of $\Phi_{\arg\min Y_i}$ is better than that of p, and the module will select $\Phi_{\arg\min Y_i}$. Otherwise, the module will select the pretrained heartbeat signal p. This selection strategy can adaptively select the best-quality heartbeat signal as the enhanced pseudo-label. The detailed training procedures of stage two are presented in Algorithm 1.

Experiments

Datasets and Implementation

Equipleth Dataset. The public radar dataset (Vilesov et al. 2022) contains 550 paired facial video and FMCW radar recordings from 91 subjects. Skin tones are categorized by the Fitzpatrick scale (Sachdeva 2009): 28 light, 49 medium, and 14 dark for skin tone fairness evaluation. Each subject contributed six 30-second recordings. More details are presented in the supplementary materials.

RHB Dataset. Our collected RHB dataset comprises 240 instances of FMCW radar data from 80 volunteers. Each participant has three separate 30-second recording sessions while maintaining a seated position approximately 0.5 to 1 meter in front of the radar acquisition board, consistent with (Vilesov et al. 2022). Data was captured at 120 frames per second. The detailed dataset configuration is presented in the supplementary materials.

Implementation Details. Following prior work (Vilesov et al. 2022), we use 10-second windows for training and heart rate evaluation. For the Equipleth dataset, we use the same training protocol as (Vilesov et al. 2022). For our RHR

Method	Equipleth			RHB→Equipleth			RHB			Equipleth→ RHB		
	MAE↓	RMSE↓	r↑	MAE↓	RMSE↓	r↑	MAE↓	RMSE↓	r↑	MAE↓	RMSE↓	r↑
FFT-based RF (Alizadeh et al. 2019)▲	13.51	21.07	0.24	13.51	21.07	0.24	12.25	18.37	0.26	12.25	18.37	0.26
(Tu, Hwang, and Lin 2016)▲	5.50	11.68	0.64	5.50	11.68	0.64	-	-	-	-	-	-
Equipleth RF (Vilesov et al. 2022)♦	**2.18**	**6.12**	**0.89**	4.53	9.63	0.65	**3.19**	**7.18**	**0.82**	2.68	6.29	0.86
mmFormer (Hu et al. 2024)♦	6.50	11.10	0.52	7.72	11.73	0.40	8.89	12.77	0.28	7.00	7.83	0.47
VitaNet (Khan et al. 2022)♦	3.14	7.70	0.77	7.43	11.86	0.40	5.28	9.25	0.66	**2.38**	**5.14**	**0.90**
Radar-APLANC (ours)★	3.95	9.72	0.64	**4.10**	**8.51**	**0.72**	3.92	7.94	0.77	3.52	7.45	0.79

▲: Traditional Training-free Methods, ♦: Supervised Methods, ★: Unsupervised Methods

Table 1: Intra-dataset and cross-dataset heart rate results of radar modality on Equipleth dataset and our RHB dataset. The best results are in bold, and the second-best results are underlined.

Method	Modality	Performance (Fairness)		
		MAE↓(→0)(bpm)	RMSE↓(→0)(bpm)	r↑(→0)
CHROM (De Haan and Jeanne 2013) ▲	RGB	7.45(4.97)	13.38(4.17)	0.46(-0.38)
ICA (Poh, McDuff, and Picard 2010)▲		8.38(4.42)	14.03(3.15)	0.41(-0.36)
PhysNet (Yu, Li, and Zhao 2019) ♦		1.78(2.22)	5.26(4.05)	0.91(-0.25)
FusionPhys-RGB (Ying et al. 2025) ♦		1.49(1.23)	5.53(3.50)	0.89(-0.12)
FFT-based RF (Alizadeh et al. 2019)▲	Radar	13.51(2.25)	21.07(2.47)	0.240(-0.25)
VitaNet (Khan et al. 2022)♦		3.14(**0.30**)	7.70(**0.48**)	0.77(**-0.04**)
mmFormer (Hu et al. 2024) ♦		6.50(0.34)	11.10(3.13)	0.52(-0.28)
Equipleth RF (Vilesov et al. 2022) ♦		2.18(0.51)	6.12(0.85)	0.89(-0.13)
Radar-APLANC (ours)★		3.95(0.91)	9.72(0.98)	0.64(-0.06)

▲: Traditional Training-free Methods, ♦ Supervised Methods, ★: Unsupervised Methods

Table 2: Heart rate performance and skin tone fairness on Equipleth dataset among radar methods and RGB methods. Among the radar methods, the best results are in bold, and the second-best results are underlined.

dataset, we apply 4-fold cross-validation. Each fold contains 50 subjects for training, 10 for validation, and 20 for testing. Both the noise and heartbeat extractors are randomly initialized and trained in two stages. The augmented pseudo-label generator inherits pretrained models from stage one. Each phase is trained with AdamW (learning rate 1e-4) for 200 epochs, and the best epoch is chosen by validation sets. The heart rate evaluation follows prior work using mean absolute error (MAE), root mean squared error (RMSE), and Pearson correlation (r).

Comparisons with State-of-the-art Methods

Intra-dataset Testing. Table 1 presents the performance of our unsupervised method on the Equipleth and our proposed RHB datasets. On the Equipleth dataset, supervised methods lead—Equipleth RF method (Vilesov et al. 2022) and VitaNet (Khan, Rigazio, and Shahzad 2022) rank first and second, respectively. Impressively, our unsupervised method achieves comparable accuracy, with only 25.8% higher MAE than VitaNet. On RHB dataset, our approach achieves near state-of-the-art performance (MAE=3.92), just 22.9% above the Equipleth RF method. Notably, our method shows strong stability on the two datasets, whereas other supervised models degrade significantly on RHB dataset. In addition, our unsupervised method also significantly outperforms traditional training-free radar methods. These results highlight the stability and robustness of Radar-APLANC.

Cross-dataset Testing. Our unsupervised method demonstrates strong generalization across datasets. When trained on RHB and tested on Equipleth (RHB→Equipleth), it incurs only a +3.8% MAE increase, significantly outperforming supervised methods (Equipleth Radar: +107.8%, VitaNet: +136.6%). In the reverse setting (Equipleth→RHB), all methods improve, yet our method demonstrates superior stability, exhibiting minimal performance deviation between direct training-testing on RHB and cross-dataset transfer from Equipleth to RHB. Notably, our MAE fluctuation across intra- and cross-dataset evaluations remains within 0.4 bpm. These results validate the robustness of our framework in learning domain-invariant representations, while supervised approaches exhibit unstable cross-domain transfers.

Performance and Fairness Comparison with Other Modalities

Table 2 compares the performance and fairness across sensing modalities in the EquiPleth dataset. Fairness is the per-

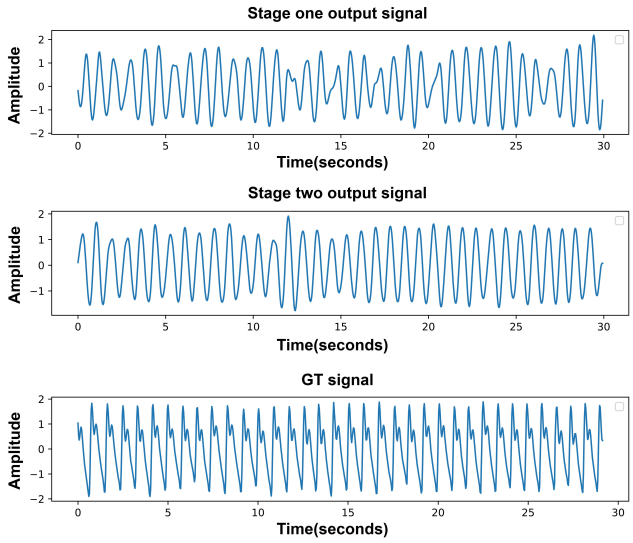

Figure 4: Example heart pulse signals generated by our stage one and stage two models and the ground truth signal.

Stage 1		Stage 2		Metrics		
Noise Matrix	Pseudo label	Augumented Pseudo label	Noise Matrix	MAE↓	RMSE↓	r↑
✓				34.48	38.34	0.01
	✓			8.94	15.88	0.30
✓	✓			4.40	9.89	0.63
✓	✓	✓		7.42	13.61	0.38
✓	✓	✓	✓	**3.95**	**9.72**	**0.64**

Table 3: Ablation study of the critical components in two stages on EquiPleth dataset. The best result is in bold.

Pretrained Heartbeat Signal	Traditional Heartbeat Signals	Noise Signal	MAE↓	RMSE↓	r↑
✓			4.56	10.09	0.63
✓	✓		8.75	15.09	0.30
	✓	✓	14.48	17.76	0.13
✓	✓	✓	**3.95**	**9.72**	**0.64**

Table 4: Ablation study on Augmented Pseudo Label Generator on EquiPleth dataset. The best results are in bold.

formance difference between dark and light skin tones, and being close to zero indicates high fairness. RGB-based methods—especially supervised ones (e.g., PhysNet (Yu, Li, and Zhao 2019))—achieve high overall performance but exhibit low fairness, indicating significant performance drops on dark skin tones due to low reflected light from dark skin. In contrast, radar-based approaches generally demonstrate better fairness than RGB since radar is independent of skin tones and lighting conditions. These findings align with prior work (Vilesov et al. 2022) that RGB methods are more prone to skin tone bias than radar-based ones. Notably, our unsupervised radar method exhibits comparable fairness with other radar methods and outperforms RGB methods in terms of fairness, highlighting the potential of label-free radar-based sensing for equitable heartbeat monitoring.

Ablation Study and Visualization

Table 3 presents the ablation study results of our method on the Equipleth dataset. When only pseudo-labels are utilized in stage one, our approach achieves an MAE of 8.94, surpassing the traditional FFT-based RF method (MAE 13.51) by a significant margin (Alizadeh et al. 2019). While the noise matrix alone fails to converge (MAE 34.48), its combination with pseudo-labels substantially reduces the MAE to 4.4, less than half of the 8.94 using only pseudo-labels. This demonstrates the feasibility of effectively leveraging both noise matrices and pseudo-labels. Furthermore, our augmented pseudo-label approach in stage two consistently outperforms all configurations in stage one. Specifically, when all configurations in both stages are enabled, it achieves the lowest MAE of 3.95. These results collectively validate the efficacy of our augmented pseudo-labeling strategy and noise matrices.

Table 4 presents the ablation study of the augmented pseudo-label generator in stage two. It can be observed that traditional heartbeat signals combined with the pretrained heartbeat signal or the noise signal for signal quality assessment and decision-making yield poor performance (MAE 8.75 and 14.48). Although directly employing the pretrained heartbeat signal as pseudo labels yields better results than the former two, it (MAE 4.56) falls short compared to the first-stage performance (MAE 4.40) shown in Table 3, thereby diminishing the value of label augmentation. By combining all three types of signals, the best performance is achieved.

As shown in Figure 4, stage one already produces heartbeat signals closely aligned with the ground truth signal, with only minor deviations in some segments. Stage two further refines these outputs, effectively reducing residual fluctuations and achieving improved waveform accuracy. The visual result in stage one demonstrates the effectiveness of NCT loss and pseudo-label in the radar unsupervised learning. In addition, the visual result in stage two reveals that augmented pseudo-labels can further improve heartbeat signals. The visualization results are also consistent with the ablation study in Table 3.

Conclusion

This paper presents Radar-APLANC, the first unsupervised framework for radar-based heartbeat sensing. To eliminate reliance on physiological signal labels and improve noise robustness, we introduce a novel NCT loss that contrasts predicted heartbeat signals with pseudo-labels and noise signals. We further design an augmented pseudo-label generator to improve the pseudo-label quality to further improve the unsupervised learning performance. Unsupervised Radar-APLANC is comprehensively validated to be comparable with other supervised methods on the public Equipleth dataset and our collected RHB dataset.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant No. 62572249), the Natural Science Foundation of Jiangsu Province (Grant No. BK20250742), and the Startup Foundation for Introducing Talent of NUIST (Grant No. 1083142501006).

References

Alizadeh, M.; Shaker, G.; Almeida, J. C. M. D.; Morita, P. P.; and Safavi-Naeini, S. 2019. Remote Monitoring of Human Vital Signs Using mm-Wave FMCW Radar. *IEEE Access*, 7: 54958–54968.

Chen, C.; Xie, W.; Huang, W.; Rong, Y.; Ding, X.; Huang, Y.; Xu, T.; and Huang, J. 2019. Progressive feature alignment for unsupervised domain adaptation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 627–636.

Choi, J.; Jeong, M.; Kim, T.; and Kim, C. 2019. Pseudo-Labeling Curriculum for Unsupervised Domain Adaptation. In *Proceedings of the British Machine Vision Conference*, 67. Cardiff, UK.

De Haan, G.; and Jeanne, V. 2013. Robust pulse rate from chrominance-based rPPG. *IEEE transactions on biomedical engineering*, 60(10): 2878–2886.

Diamant, I.; Rosenfeld, A.; Achituve, I.; Goldberger, J.; and Netzer, A. 2024. De-confusing pseudo-labels in source-free domain adaptation. In *European Conference on Computer Vision*, 108–125. Springer.

Droitcour, A. D. 2006. *Non-contact measurement of heart and respiration rates with a single-chip microwave doppler radar*. Stanford University.

Gideon, J.; and Stent, S. 2021. The Way to My Heart Is Through Contrastive Learning: Remote Photoplethysmography From Unlabelled Video. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 3995–4004.

Hu, Q.; Zhang, Q.; Lu, H.; Wu, S.; Zhou, Y.; Huang, Q.; Chen, H.; Chen, Y.-C.; and Zhao, N. 2024. Contactless Arterial Blood Pressure Waveform Monitoring with mmWave Radar. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 8(4): 1–29.

Khan, U. M.; Rigazio, L.; and Shahzad, M. 2022. Contactless Monitoring of PPG Using Radar. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 6(3).

Lai, Z.; Vesdapunt, N.; Zhou, N.; Wu, J.; Huynh, C. P.; Li, X.; Fu, K. K.; and Chuah, C.-N. 2023. Padclip: Pseudo-labeling with adaptive debiasing in clip for unsupervised domain adaptation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 16155–16165.

Li, Z.; and Yin, L. 2023. Contactless pulse estimation leveraging pseudo labels and self-supervision. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 20588–20597.

Lin, Y.; Xie, L.; Wu, Y.; Yan, C.; and Tian, Q. 2020. Unsupervised Person Re-Identification via Softened Similarity Learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.

Litrico, M.; Del Bue, A.; and Morerio, P. 2023. Guiding pseudo-labels with uncertainty estimation for source-free unsupervised domain adaptation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 7640–7650.

Mercuri, M.; Lorato, I. R.; Liu, Y.-H.; Wieringa, F.; Hoof, C. V.; and Torfs, T. 2019. Vital-sign monitoring and spatial tracking of multiple people using a contactless radar-based sensor. *Nature Electronics*, 2(6): 252–262.

Morerio, P.; Volpi, R.; Ragonesi, R.; and Murino, V. 2020. Generative pseudo-label refinement for unsupervised domain adaptation. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 3130–3139.

Poh, M.-Z.; McDuff, D. J.; and Picard, R. W. 2010. Advancements in noncontact, multiparameter physiological measurements using a webcam. *IEEE transactions on biomedical engineering*, 58(1): 7–11.

Sachdeva, S. 2009. Fitzpatrick skin typing: Applications in dermatology. *Indian journal of dermatology, venereology and leprology*, 75: 93.

Saito, K.; Ushiku, Y.; and Harada, T. 2017. Asymmetric tri-training for unsupervised domain adaptation. In *International conference on machine learning*, 2988–2997. PMLR.

Seo, S.; Lee, J.-Y.; and Han, B. 2022. Unsupervised learning of debiased representations with pseudo-attributes. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 16742–16751.

Shi, J.; Yin, X.; Chen, Y.; Zhang, Y.; Zhang, Z.; Xie, Y.; and Qu, Y. 2024. Multi-memory matching for unsupervised visible-infrared person re-identification. In *European Conference on Computer Vision*, 456–474. Springer.

Song, R.; Lu, Z.; Zhang, D.; Fang, L.; Wu, Z.; Hu, Y.; Sun, Q.; and Chen, Y. 2024. Unleashing the potential of self-supervised rf learning with group shuffle. *IEEE Transactions on Mobile Computing*.

Song, R.; Zhang, D.; Wu, Z.; Yu, C.; Xie, C.; Yang, S.; Hu, Y.; and Chen, Y. 2022. Rf-url: unsupervised representation learning for rf sensing. In *Proceedings of the 28th Annual International Conference on Mobile Computing And Networking*, 282–295.

Speth, J.; Vance, N.; Flynn, P.; and Czajka, A. 2023. Non-contrastive unsupervised learning of physiological signals from video. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 14464–14474.

Sun, Z.; Chen, S.; Yao, T.; Yin, B.; Yi, R.; Ding, S.; and Ma, L. 2023. Contrastive pseudo learning for open-world deepfake attribution. In *Proceedings of the IEEE/CVF international conference on computer vision*, 20882–20892.

Sun, Z.; and Li, X. 2022. Contrast-phys: Unsupervised video-based remote physiological measurement via spatiotemporal contrast. In *European Conference on Computer Vision*, 492–510. Springer.

Tu, J.; Hwang, T.; and Lin, J. 2016. Respiration Rate Measurement Under 1-D Body Motion Using Single

Continuous-Wave Doppler Radar Vital Sign Detection System. *IEEE Transactions on Microwave Theory and Techniques*, 64(6): 1937–1946.

Vilesov, A.; Chari, P.; Armouti, A.; Harish, A. B.; Kulkarni, K.; Deoghare, A.; Jalilian, L.; and Kadambi, A. 2022. Blending camera and 77 GHz radar sensing for equitable, robust plethysmography. *ACM Trans. Graph.*, 41(4): 36–1.

Wang, K.; Thakur, N.; Reimers, N.; and Gurevych, I. 2022. GPL: Generative Pseudo Labeling for Unsupervised Domain Adaptation of Dense Retrieval. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT)*, 2345–2360. Seattle, USA: Association for Computational Linguistics.

Wu, Z.; Xie, Y.; Zhao, B.; He, J.; Luo, F.; Deng, N.; and Yu, Z. 2025. CardiacMamba: A Multimodal RGB-RF Fusion Framework with State Space Models for Remote Physiological Measurement. *IEEE Transactions on Instrumentation and Measurement*.

Ying, C.; Yang, H.; Ge, J.; Sun, Z.; Cheng, X.; Ren, K.; and Li, X. 2025. FusionPhys: A Flexible Framework for Fusing Complementary Sensing Modalities in Remote Physiological Measurement. In *International Conference on Computer Vision*.

Yu, Z.; Li, X.; and Zhao, G. 2019. Remote photoplethysmograph signal measurement from facial videos using spatio-temporal networks. *British Machine Vision Conference*.

Yue, Z.; Shi, M.; and Ding, S. 2023. Facial video-based remote physiological measurement via self-supervised learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(11): 13844–13859.

Zhang, H.; Zhang, D.; Song, R.; Wu, Z.; Chen, J.; Fang, L.; Lu, Z.; Hu, Y.; Lin, H.; and Chen, Y. 2025. Umimo: Universal unsupervised learning for mmwave radar sensing with mimo array synthesis. *IEEE Transactions on Mobile Computing*.