Non-Contact Sleep Apnea Detection via Subtle Motion Analysis of UWB Radar

Xin Liu[1], Xinyu Li[1], Yutong Guo[1], Shi Fang[2], Xinke Wang[3],
Wei Qian[1], Fusang Zhang[4], Jinyang Huang[1*]
[1] Hefei University of Technology, [2] University of Science and Technology of China,
[3] Anhui University, [4] Beihang University.

Subtle physiological motions, such as respiration-induced chest displacement, carry rich cues for sleep apnea assessment, yet remain largely imperceptible to conventional visual imaging systems. Although clinical polysomnography can accurately capture subtle motion changes, it is not suitable for long-term home monitoring. To address this issue, we propose a Non-contact Sleep Apnea Detection (NcSAD) based on ultra wideband Radar. Specifically, we develop a robust signal processing pipeline to suppress radar noise, including IQ signal adaptive rotation for precise phase recovery, respiratory signal selection with signal-to-noise ratio, outlier filtering, and waveform correction. Then, we introduce a self-supervised contrastive learning paradigm to eliminate the reliance on label. To avoid noise that may be introduced by the random augmentation function for time-series signals, we further design an augmentation strategy based on singular spectrum analysis that generates semantically consistent positive samples via multi-scale signal decomposition and reconstruction. To validate the method, we release a Radar Obstructive Sleep Apnea (ROSA) dataset consisting of 10,203 samples of 30 subjects. Extensive experiments on ROSA show outstanding performance, offering a promising solution for active self-assessment and early detection of sleep disorders.

1. Introduction

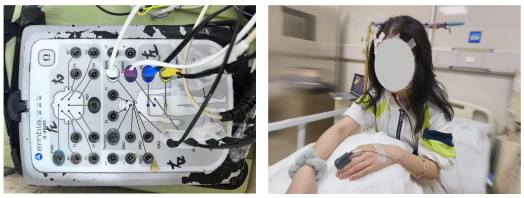
Accurate motion recognition requires the capture of subtle movements that encode fine-grained dynamics beyond coarse motion patterns [9, 18]. During sleep, the human body exhibits subtle and continuous motions, most notably the slight rise and fall of the chest caused by breathing. These movements are closely related to breathing activity and can provide useful cues for monitoring sleep-related conditions. In principle, they can be observed as continuous temporal changes, where normal breathing tends to produce stable and repetitive patterns, while abnormal events often introduce sleep apnea.

However, such subtle motions are difficult to reliably capture using conventional visual imaging systems. The motion amplitude is often very small, and can be easily affected by lighting conditions, occlusions, and slight posture variations during sleep. As a result, important breathing-related motions are not always clearly visible in standard visual observations, particularly in long-term or unconstrained monitoring scenarios. These challenges are particularly relevant for sleep apnea detection. In central sleep apnea (CSA), breathing cessation is typically accompanied by the absence of chest movement, while in obstructive sleep apnea (OSA), chest movement may persist even when airflow is blocked [15, 25]. As a result, similar motion patterns may correspond to different underlying conditions, making it difficult to distinguish between them based on motion information alone. This highlights the need for approaches that can better utilize subtle motion cues while handling noise and ambiguity in real-world sleep environments.

Clinically, CSA occurs at relatively low frequencies [12, 26], while OSA is substantially more common yet more challenging to detect due to the presence of confounding chest motion during apnea. Epidemiological studies estimate that nearly one billion adults aged 30 to 69 worldwide suffer from OSA, with 425 million experiencing moderate-to-severe symptoms [4, 24]. Moreover, the increasing prevalence of underlying health conditions (*e.g.*, obesity and hypertension) continues to exacerbate the global burden of sleep apnea [17]. Currently, polysomnography (PSG) monitoring is the universal standard for sleep apnea diagnosis [1, 16, 21]. However, as shown in Fig. 1, PSG is not suitable for daily or home scenarios due to the requirement for professional medical personnel to wear PSG leads on subjects, the uncomfortable feelings of the invasive PSG device, and the high price of PSG. This issue makes PSG monitoring unable to achieve early detection of sleep apnea, which inevitably delays timely intervention and treatment [6, 30]. Therefore, developing a non-contact sleep apnea detection method is valuable and a general trend.

Recently, wireless sensing, particularly millimeter-wave radar, has emerged as a promising solution due to its ability to capture subtle chest movements associated with

*Corresponding author: hjy@hfut.edu.cn



(a) PSG (b) PSG electrode patch

Figure 1. Clinical data collection device and scenario.

respiration [7, 37]. Existing radar-based apnea detection approaches typically fall into two categories, *i.e.*, traditional machine learning-based methods, and deep learning-based approaches. Traditional machine learning methods [13, 27, 28, 35] extract hand-crafted spectral features and train classical classifiers, whereas deep learning approaches [14, 20, 21] leverage neural networks to automatically learn discriminative representations for sleep apnea detection.

Despite rapid progress, non-contact sleep apnea monitoring remains constrained by three fundamental challenges: ***(1) Lack of Publicly Available Non-contact Sleep Apnea Datasets:*** Most existing radar-based apnea studies collect small-scale or private datasets. The absence of open benchmarks significantly hinders the development of non-contact sleep apnea research. ***(2) Robust Extraction of Signals in Noisy Environments:*** Radar measurements during sleep are easily contaminated by diverse sources of interference, *e.g.*, environmental clutter, background noise, and involuntary body movements. Existing methods seldom address these complex noise conditions comprehensively, making reliable respiration reconstruction a major challenge. ***(3) The High Cost of Manual Annotation:*** PSG monitoring only automatically labels a portion of sleep apnea events, requiring doctors to perform additional annotations based on PSG data, which is costly. To avoid reliance on labels, contrastive learning is a well-suited paradigm, however, existing random perturbation augmentation functions for time-series data are not entirely applicable to respiratory waveforms due to the potential introduction of additional noise.

To overcome the above challenges, we collect and release a radar-based obstructive sleep apnea (ROSA) dataset that contains 10203 one-minute respiratory samples from 30 radar-monitored subjects and PSG monitoring data from 304 subjects. Upon this, we propose non-contact sleep apnea detection (NcSAD) method, as shown in Fig. 2. Specifically, We designed a denoising module to suppress noise in raw radar data and breathing waveforms, including signal adaptive selection to avoid phase wrapping, adaptive selection of amplitude or phase sequences as respiratory signals based on signal-to-noise ratio (SNR), and outlier removal and correction in respiratory waveforms, to improve sample quality. In practice, such noise is ubiquitous in radar signals. Therefore, the proposed noise suppression and preprocessing is a necessary component for robust radar signal modeling. Importantly, this module is designed as a general radar-based physiological sensing solution, not restricted to the specific dataset used in this study, demonstrating its potential for broader applicability across different devices and environments. Then, to avoid dependence on labels during training, we introduce a self-supervised contrastive learning paradigm. Furthermore, we design an augmentation function based on singular spectrum analysis (SSA), which utilizes multi-scale decomposition and reconstruction of the signal to ensure the semantic consistency of the generated respiratory samples. Extensive experiments on the ROSA dataset highlight the advantages of NcSAD.

In brief, our contributions can be summarized as follows:

- We release a novel radar-based dataset of obstructive sleep apnea (ROSA) and propose a non-contact sleep apnea monitoring method based on contrastive learning.
- To suppress radar noise, we propose a series of denoising strategies for radar original data and respiratory waveforms to improve data quality, which benefits both downstream classification and the robustness of representations extracted from subtle respiratory variations.
- We introduce contrastive learning to avoid label dependency, and design an SSA-based augmentation function to generate semantically consistent breathing signals.
- Extensive experiments on ROSA dataset demonstrate that NcSAD achieves superior performance in sleep apnea detection. Furthermore, correlation experiments show a significant correlation (0.91) between NcSAD and PSG monitoring results, proving the reliability of the method.

2. Related Work

2.1. Sleep Apnea Detection

The evolution of sleep apnea detection has transitioned from traditional clinical approaches to intelligent sensing technologies. Early practices relied on polysomnography [2, 36]. Despite accurate, these contact-based methods suffer from cumbersome equipment and high costs, making them unsuitable for long-term daily monitoring due to its equipment complexity. Recently, non-contact sensing based on UWB radar has emerged as a promising alternative. Researchers have developed energy-efficient frameworks [4] and attention-enhanced models [21] to capture respiratory features remotely.

However, a critical limitation persists in these non-contact approaches due to their reliance on supervised training, which requires large-scale expert annotation and is costly and time-consuming. Even existing works attempting to utilize contrastive learning often lack specific augmentation strategies tailored for radar data, limiting their robustness against real-world noise. To overcome these dual challenges of physical intrusiveness and annotation de-

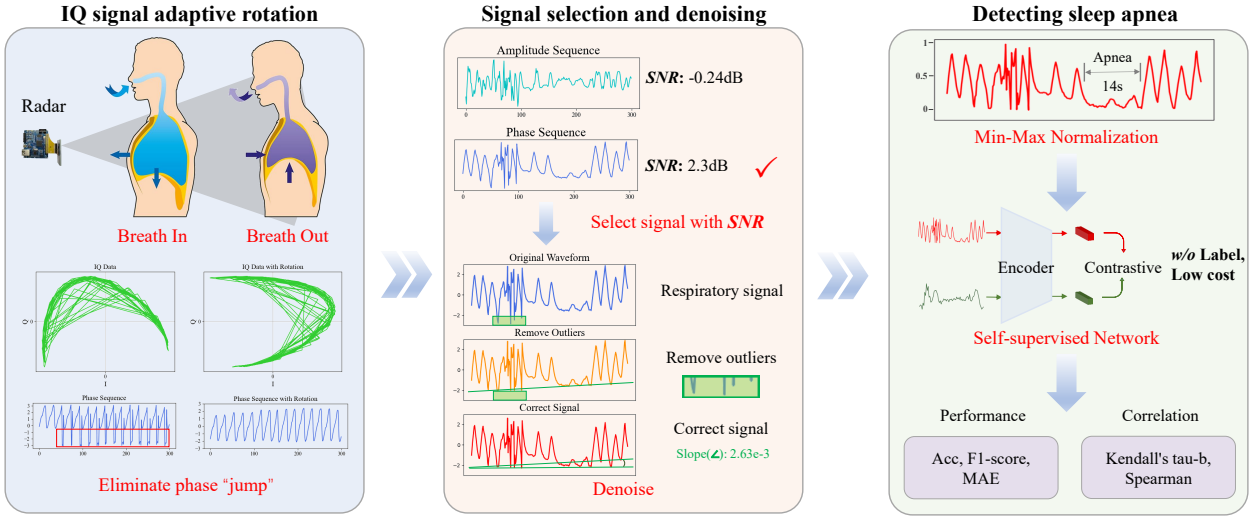


Figure 2. Overall framework of non-contact sleep apnea detection, including IQ signal adaptive rotation, signal selection and denoising, and sleep apnea detection.

pendence, we employ a Transformer-based self-supervised contrastive learning paradigm to avoid relying on labels during the training process and design a novel SSA-based augmentation function to generate enhanced samples with semantic consistency and diversity.

2.2. Ultra-Wideband Radar Perception

Ultra-Wideband (UWB) radar technology has established a systematic technical evolution pathway in the field of non-contact physiological signal sensing. Li *et al.* [20] proposed an attention-enhanced UWB signal-based respiratory monitoring system for bed positioning and sleep apnea detection. By introducing a multi-window voting mechanism, they further optimized the algorithm, achieving significant progress in nocturnal fine-grained respiratory monitoring [19]. Subsequently, Li *et al.* [22] extended their previous work and proposed the Hypnos system, which achieved precise classification of five sleep stages by incorporating sleep apnea detection as an auxiliary task.

In other sensing tasks, *i.e.*, cardiac activity monitoring, Wang *et al.* [33] proposed utilizing UWB signals to generate Electrocardiogram (ECG)-level cardiac waveforms. For physiological signal sensing in dynamic scenarios, Han *et al.* [10] converted UWB signals into two-dimensional thermal maps and extracted respiratory motion features. Meanwhile, Zheng *et al.* [39] achieved respiratory waveform reconstruction and trajectory tracking of walking subjects through radar pair cooperative localization and an In-phase and Quadrature Variational Encoder-Decoder architecture. However, despite these significant advances, the application of UWB radar-based sensing technology in long-term sleep apnea monitoring still lacks noise suppression of the raw radar signal and respiratory waveform.

3. Radar Data Noise Suppression

3.1. IQ Signal Adaptive Rotation

The signals collected by the UWB radar consist of in-phase (I) and quadrature (Q) components. Accordingly, the received echo signal $S(t_i)$ at the sampling instant t_i can be represented as:

$$S(t_i) = I(t_i) + jQ(t_i), \quad i = 1, 2, \ldots, N. \tag{1}$$

To further extract physically meaningful characteristics, a complex-to-polar transformation is applied to the IQ signal, the corresponding amplitude and phase sequences can be represented as follows:

$$A(t_i) = \sqrt{I^2(t_i) + Q^2(t_i)}, \quad \phi(t_i) = \arctan \frac{Q(t_i)}{I(t_i)}, \tag{2}$$

where the amplitude $A(t_i)$ characterizes the strength of the received radar echo, while the phase $\phi(t_i)$ represents the propagation delay caused by chest movement.

Since the phase component is highly sensitive to subtle movements, it enables more precise and robust characterization of fine-grained human body motion, making it particularly suitable for detecting subtle respiratory variations. However, since the phase is confined to the interval $[-\pi, \pi]$, it is susceptible to phase wrapping. In particular, when the phase change between two adjacent sampling points spans π, discontinuities caused by phase wrapping inevitably occur. Therefore, it is necessary to suppress the wrapping effect between consecutive phase frames.

A common manner to solve phase wrappings is through the phase unwrapping function. However, this strategy has the obvious disadvantage of accumulating errors. If an error occurs in the unwrapping of a certain point in the process of accumulating phase differences point by point, this error will be transmitted and accumulated in the subsequent phase values, which can result in a global deviation. In particular, phase signals are easily interfered with by noise, which causes abnormal phase differences between adjacent points and is more likely to cause cumulative errors.

To accurately achieve phase unwrapping, we design an IQ signal adaptive rotation strategy on the original radar IQ data to alleviate this phenomenon. Specifically, given the signal $S(t_i) = I(t_i)+jQ(t_i)$ $(i = 1, \ldots, N)$, we first quantify its spatial distribution in the complex plane by counting the number of samples located in each quadrant. The resulting quadrant counts $\{Q_1^c, Q_2^c, Q_3^c, Q_4^c\}$ are then sorted in descending order to identify the dominant phase distribution pattern. Upon this, we adaptively rotate the IQ data in the complex plane based on the sorting results as follows:

- If Q_2^c and Q_3^c are ranked first and second, we rotate the IQ data by π.
- If Q_1^c and Q_2^c are ranked first and second, we rotate the IQ data by -0.5π.
- If Q_3^c and Q_4^c are ranked first and second, we rotate the IQ data by 0.5π.
- If Q_2^c and Q_4^c are ranked first and second, and Q_3^c is ranked third, we rotate the IQ data by π.
- If Q_1^c and Q_3^c are ranked first and second, and Q_2^c is ranked third, we rotate the IQ data by π.
- In other cases, we do not rotate the IQ data.

Since phase information captures relative motion changes, rotating the IQ signal does not destroy the validity of the respiratory waveform and restores the true phase relationship under the impacts of motions. Compared with the simple phase unwrapping methods, our strategy solves the phase wrapping more fundamentally and thoroughly from the raw radar data, while preventing phase accumulation.

3.2. Respiratory Signal Selection

After obtaining amplitude and phase sequences, the key issue is how to choose the amplitude or phase sequences as a suitable signal for further processing. In clinical practice, the degree of weakened breathing movements during sleep apnea varies among individuals, resulting in differences in the expression ability of amplitude and phase information. Hence, we design an automatic strategy to dynamically select the suitable amplitude or phase sequence as the respiratory signal for each sample.

Specifically, the respiratory frequency of normal people and patients with sleep apnea can be empirically assumed to be between 0-4 Hz, therefore, the components within which are treated as valid signals, while the remainder are regarded as noise signals. Based on this, the signal-to-noise ratio (SNR) of amplitude and phase sequences can be calculated, respectively. To ensure robust and high-quality respiratory reconstruction, the sequence with higher SNR is selected as the final respiratory waveform.

Nevertheless, unconscious movements of the subject during sleep can incur various noises, including outliers and abnormal fluctuations. Apart from this, several phase wrappings that still exist in some samples can also cause outliers. These noises seriously hinder the model from detecting sleep apnea. To address these problems, two schemes are further put forward to remove outliers and further correct the signal.

3.3. Removing Outliers

With regard to outliers, the interquartile range (IQR) is introduced to identify outliers since IRQ focuses on the middle 50% of the data and is less affected by extreme values, making it a robust measure of variability. Concretely, IQR is calculated by subtracting the first quartile (Q_1) from the third quartile (Q_3): $IQR = Q_3 - Q_1$.

Samples lying outside the valid interval $[B_{low}, B_{up}]$ are regarded as outliers, where the lower and upper bounds are determined by:

$$\begin{cases} B_{low} = Q_1 - \lambda \cdot IQR, \\ B_{up} = Q_3 + \lambda \cdot IQR, \end{cases} \tag{3}$$

where λ is a tunable hyperparameter controlling the sensitivity of outlier detection. In this work, λ is set to 0.5.

Since OSA weakens chest movement, the increase in the amplitude of sleep apnea caused by subtle unconscious movements is not a characteristic of sleep apnea, which does not affect the model's recognition of sleep apnea. Therefore, to ensure the integrity of the waveform, we only remove outliers below B_{low} and replace them with B_{low}.

3.4. Correcting Signal

To further alleviate abnormal waveform fluctuations induced by involuntary body movements during sleep, we design a slope-based correction strategy. Given a respiratory time series (waveform) $x_i \in R^N$, where R represents the set of real numbers and N denotes the number of temporal frames, we search the minimum amplitude value at the beginning ($a = min(x_i[0 : 50])$ and end($b = min(x_i[end - 50 : end])$ of the time series, and then identify the corresponding indexes (Idx_a and Idx_b) based on these minimum values. Then, the slope S of the waveform can be calculated as:

$$S = \frac{b - a}{Idx_b - Idx_a}. \tag{4}$$

Based on the calculated slope S, the position of each element in the time series can be corrected by adjusting the

waveform frame by frame to ensure that the corrected signal is as horizontal as possible.

3.5. Normalization

To increase data diversity, the angle between the radar antenna and the chest is adjusted to a maximum recognizable angle of more than 60 degrees to further collect the radar data from 13 people among all subjects. In this case, the energy of the received echo will weaken significantly, causing the signal fluctuation range to decrease. Moreover, the fluctuation ranges of the amplitude sequence and the phase sequence are different, which is not conducive to deep learning training. To eliminate these negative impacts, we normalize the denoised data based on Min-Max scaling [34] to ensure the numerical stability of the neural network and the robustness of our model to the radar placement angle.

4. Methodology

4.1. Data Augmentation

Generally, sleep monitoring data are usually collected by PSG, where partial sleep apnea events are automatically annotated by PSG while the remaining data need to be manually annotated by doctors. However, the cost of manual labeling is expensive, especially when the number of patients increases. To this end, we introduce a self-supervised contrastive learning to eliminate the limitation on label requirements during the training phase.

It is worth noting that contrastive learning usually provides self-supervised information by generating positive and negative pairs based on augmentation functions to guide model training [3, 5, 23, 32]. Therefore, data augmentation technology is extremely important and necessary for generating positive and negative samples. Unfortunately, existing augmentation functions for time-series data employ two general strategies: jitter scaling and permutation jitter. Neither of these strategies is universally applicable to sleep apnea signals. In particular, permutation jitter can damage the inherent semantic information of samples containing sleep apnea events, making it difficult for the model to learn accurate features of these events. Therefore, to ensure the quality of augmented samples, it is necessary to design a specific augmentation function for sleep apnea signals.

Inspired by [8, 31], we find that good positive samples possess two properties: (1) These positive samples and original samples have high semantic consistency, which effectively and accurately provides self-supervised information for model training. (2) These positive samples maintain a certain difference from the original samples, which is beneficial in improving the generalization of the model. With regard to this, considering that SSA can adaptively decompose nonlinear signals and realize multi-scale reconstruction to enhance the diversity of samples, which well satisfies the above requirements, we utilize SSA to obtain the spectral components of the waveform signal[11, 38], and then generate positive samples by reconstructing the waveform signal with partial spectral components.

Concretely, given a set of sequences $X = \{x_1, \ldots, x_i, \ldots x_M\}$, where M is the number of sequences. For the i-th sequence x_i, we transfer $x_i \in R^N$ (N is the total number of frames of vector x_i) into multiply series $T_1, \ldots, T_k$ ($k = N - L + 1$, L is the window length), which are called lagged vectors ($2 \leq L \leq T$). The trajectory matrix $\boldsymbol{T}$ of x_i is composed of these lagged vectors: $\boldsymbol{T} = [T_1, \ldots, T_k] = (t_{ij})_{i,j=1}^{L,K}$.

Then, let $\lambda_1, \ldots, \lambda_L$ denote the eigenvalues of $\boldsymbol{T}\boldsymbol{T}^T$ in decreasing order of magnitude, and $U_1, \ldots, U_L$ represent the standard orthogonal vectors corresponding to these eigenvalues, respectively. When $V_i = \boldsymbol{T}^T U_i / \sqrt{\lambda_i}$, we can perform singular value decomposition (SVD) on the trajectory matrix $\boldsymbol{T}$, which can be written as:

$$\boldsymbol{T} = \boldsymbol{T}_1 + \cdots + \boldsymbol{T}_d, \tag{5}$$

where $\boldsymbol{T}_i = \sqrt{\lambda_i} U_i V_i^T$, and d is the rank of $\boldsymbol{T}$.

For the reconstruction phase, the grouping operation divides all elementary matrices T_i ($i \in [1, d]$) into several linearly independent groups and sums the elementary matrices within each group. The grouping procedure is called the eigentriple grouping. Then, the operation of diagonal averaging transfers each group matrix into a time series. Therefore, the signal can be reconstructed by choosing some groups, which is not only consistent with the original trend of the signal, but also preserves essential information on signal fluctuation. In such cases, adopting signal decomposition and reconstruction based on SSA as the augmentation strategy supports high-quality positive pair construction in contrastive learning.

4.2. Capturing Signal Change

Capturing the global dependencies of time series data is essential for detecting sleep apnea. To this end, the transformer encoder is adopted as the encoder to learn the representation. Specifically, since the elements of the signal are represented as discrete amplitude values in numerical space, the step of word embedding is discarded in the transformer. In addition, since self-attention mechanism does not focus on temporal information, position encoding is indispensable. Detailed architecture of the transformer encoder is recorded in [29]. Given a respiratory signal $x_i \in R^N$, the representation $\boldsymbol{H}_i \in R^{N \times D}$ (D denotes the dimension of the representation) at the last layer can be learned as follows:

$$\boldsymbol{H}_i = Multihead(\boldsymbol{Z}_i^q, \boldsymbol{Z}_i^k, \boldsymbol{Z}_i^v), \tag{6}$$

where $\boldsymbol{Z}_i^q$, $\boldsymbol{Z}_i^k$, and $\boldsymbol{Z}_i^v$ denote the x_i projection of query, key, and value in transformer encoder. For simplicity, we

do not emphasize position encoding in Eq.(6), while it is strictly enforced in our experiments. Afterwards, $\boldsymbol{H}_i$ is simplified to $\boldsymbol{h}_i \in R^D$ as the final representation via the global maximum pooling through the time dimension.

4.3. Loss Function

In the training phase, given a respiratory signal $x_i \in X$, where $X = \{x_i\}_{i=1}^{M}$ is the set of all segmented samples composed of 30 subjects, positive samples x_i^+ and x_i^{++} can be generated via the designed augmentation function to further construct the positive pair (x_i^+, x_i^{++}), while other generated samples are treated as negative samples. Based on the learned representations of positive samples $\boldsymbol{h}_i^+$ and $\boldsymbol{h}_i^{++}$, when we treat x_i^+ as the anchor, the loss function of the positive pair $(\boldsymbol{h}_i^+, \boldsymbol{h}_i^{++})$ can be written as:

$$\ell(\boldsymbol{h}_i^+, \boldsymbol{h}_i^{++}) = -\log \frac{e^{\frac{s(\boldsymbol{h}_i^+, \boldsymbol{h}_i^{++})}{\tau}}}{e^{\frac{s(\boldsymbol{h}_i^+, \boldsymbol{h}_i^{++})}{\tau}} + \sum_{j \neq i} e^{\frac{s(\boldsymbol{h}_i^+, \boldsymbol{h}_j^{+})}{\tau}} + \sum_{j \neq i} e^{\frac{s(\boldsymbol{h}_i^+, \boldsymbol{h}_j^{++})}{\tau}}}, \tag{7}$$

where $j \in [1, M]$, $s(\cdot, \cdot)$ denotes the cosine similarity, and τ is a temperature parameter.

Symmetrically, when x_i^{++} is treated as the anchor, the loss function is $\ell(\boldsymbol{h}_i^{++}, \boldsymbol{h}_i^{+})$. Therefore, the final loss function is formulated as:

$$L = \frac{1}{2M} \sum_{i=1}^{M} \ell(\boldsymbol{h}_i^+, \boldsymbol{h}_i^{++}) + \ell(\boldsymbol{h}_i^{++}, \boldsymbol{h}_i^{+}), \tag{8}$$

where M is the number of training data.

5. Implementation and Evaluation

5.1. Dataset

To promote the development of sleep apnea detection with radar, we establish a radar-based obstructive sleep apnea, which contains 10203 radar samples. All samples are clinically collected with radar AW-UWB-EV-02 in the Respiratory and Critical Care Department III of the Second Affiliated Hospital of Anhui Medical University, where the radar has the properties of single transmit and receive antennas, low frequency, low power consumption, long continuous working time. In clinical settings, the radar transmitting and receiving antennas are oriented toward the chest cavity of the subject. To ensure the naturalness of clinical data and simulate practical scenarios, we do not impose any background environment requirements and physical movement constraints on subjects, *e.g.*, covering, sleeping posture, *etc*. Until now, valid sleep apnea data from 304 subjects have been collected with PSG, including 30 subjects monitored with radar. Due to the wide range of sleep apnea indicators, only partial sleep apnea events can be automatically marked by PSG based on these indicators, while the remaining are manually annotated by the professional medical staff. Currently, most private non-contact sleep apnea datasets involve fewer than ten subjects. In practice, collecting radar-based sleep data with reliable annotations is inherently challenging in real-world scenarios. Despite this, the ROSA dataset captures realistic conditions and clinically relevant variations. Besides, the original data are also augmented to some extent through the contrastive learning augmentation function, which helps to improve generalization. Detailed statistics on ROSA are provided in Table 1.

Table 1. Subject distribution and statistics

Indicators	**Statistics**
All subjects	312
Valid subjects	304
Valid radar monitoring	30
Radar disconnected	3
PSG lead disconnectiona	5
Gender-Male	204(74.73%)
Gender-Female	69(25.27%)
Age	40.12 ± 14.25 (14-84)
BMI-Obese	160(58.61%)
Hypertension before sleep	101(40.00%)
Hypertension after sleep	151(55.31%)
Sleep apnea-Severe	153(56.04%)

Since the sleep apnea dataset is initially established for sleep apnea classification or detection tasks, the radar data are first divided into different samples. Taking into account that obstructive sleep apnea usually lasts more than 10 seconds, we adopt a one-minute sliding window and a one-minute step size to segment radar data based on PSG, which is similar to the segmentation manner used in existing work[28]. Besides, to increase the comprehensiveness of the data, we additionally set a different angle between the radar antenna and the chest cavity to a maximum recognizable angle of 60 degrees for 13 subjects. Since the sampling rate of AW-UWB-EV-02 is 2042 frames/s, we downsample the sample to 5 frames/s to reduce redundant information.

5.2. Metrics

For the two-classification task (apnea or non-apnea), accuracy (Acc) and F1-score are used as evaluation metrics. For three-classification classification tasks (number of sleep apnea events), extended versions of the F1-score, *i.e.*, Micro F1-score, Macro F1-score, and weighted F1-score, are leveraged to evaluate performance. Specifically, since Micro F1-score is identical to the accuracy rate indicator in the three-category classification, Micro F1-score is thus replaced by Acc. The Macro F1-score treats each class equally, and its value can be affected by rare categories. However, the weighted F1-score allocates different weights

Table 2. Results of two-classification tasks.

Methods	N-class	Acc	F1-score
NcSAD-BiGRU	2	88.05 ± 3.10	82.49 ± 5.09
NcSAD-BiLSTM	2	84.14 ± 4.08	67.87 ± 4.61
NcSAD-Conv1D	2	93.24 ± 0.59	90.55 ± 0.96
NcSAD	2	**97.23 ± 1.26**	**96.13 ± 1.83**
NcSAD-Supervised	2	98.05 ± 0.76	97.28 ± 1.05

Table 3. Results of three-classification tasks.

Methods	N-class	Acc	Macro F1-score	Weighted F1-score
NcSAD-BiGRU	3	86.56 ± 7.23	77.62 ± 8.88	85.52 ± 8.71
NcSAD-BiLSTM	3	79.69 ± 2.45	64.27 ± 5.68	73.09 ± 3.88
NcSAD-Conv1D	3	88.09 ± 0.91	78.24 ± 1.53	87.24 ± 1.53
NcSAD	3	**96.25 ± 1.03**	**94.50 ± 1.53**	**96.24 ± 1.01**
NcSAD-Supervised	3	97.46 ± 1.17	95.54 ± 1.78	97.42 ± 1.21

for different categories, whose weight is determined according to the true distribution ratio of the category. To unbiasedly evaluate model performance, the five-fold cross-validation is adopted to report the mean and standard deviation of all experimental results.

5.3. Overall Performance

To verify the effectiveness of our method, comparative experiments are performedSince most methods in this field are not open source, we replaced the Transformer encoder with other encoders, including Conv1D, BiLSTM, and BiGRU. We reproduce these networks and directly replaced the Transformer encoder used in our framework. The results are reported in Table 2 and Table 3.

Obviously, NcSAD achieves the best performance in both the two-classification and the three-classification tasks under self-supervised training. On the contrary, although BiLSTM and BiGRU are commonly used for processing time series data, they mainly rely on local temporal dependencies in sequence modeling, making it difficult to effectively capture the global dynamic patterns in respiratory signals. Therefore, their performance in this task is limited, with Acc of 79.69% and 86.56% on three-classification task, respectively. Furthermore, to intuitively illustrate the advantages of the proposed method, we conduct a visual analysis of the experimental results, as shown in Fig. 3. It can be observed that NcSAD produces significantly smaller detection errors compared to other methods. This can be attributed to the fact that Transformer can effectively capture the global dependencies of respiratory signals, enabling the model to accurately learn the features of sleep apnea.

In addition, due to the inherent class imbalance in sleep apnea data, the F1 scores in both classification tasks are lower than the ACC, which is a common challenge in medical dataset studies. Nevertheless, within the NcSAD, the

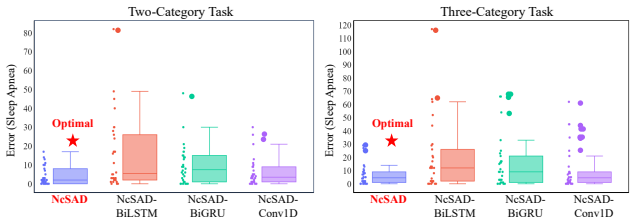


Figure 3. Evaluation errors for different models.

Table 4. Results of ablation study on three-classification tasks.

NcSAD *w/o* Denoising	2	96.02 ± 0.62	94.37 ± 0.99
NcSAD *w/o* Rotation	2	94.96 ± 1.70	92.68 ± 2.44
NcSAD *w* Pha	2	96.25 ± 1.01	94.74 ± 1.49
NcSAD *w* Amp	2	95.47 ± 1.21	93.53 ± 1.97
NcSAD *w* One Aug	2	95.23 ± 0.61	93.19 ± 1.07

Table 5. Results of ablation study on three-classification tasks.

Methods	N-class	Acc	Macro F1-score	Weighted F1-score
NcSAD *w/o* Denoising	3	95.20 ± 0.55	93.47 ± 0.76	95.17 ± 0.58
NcSAD *w/o* Rotation	3	94.77 ± 1.57	93.13 ± 1.94	94.72 ± 1.59
NcSAD *w* Pha	3	95.23 ± 0.77	93.77 ± 1.00	95.21 ± 0.78
NcSAD *w* Amp	3	95.04 ± 0.61	93.57 ± 0.78	95.01 ± 0.61
NcSAD *w* One Aug	3	94.02 ± 1.09	91.95 ± 1.61	93.99 ± 1.10

F1-score, Micro F1-score, Macro F1-score, and Weighted F1-score exhibit no significant decline, indicating that our model effectively mitigates the impact of class imbalance through contrastive learning. The underlying reason lies in the proposed contrastive learning paradigm: on one hand, it enhances the discriminative ability of minority class samples by strengthening feature differentiation; on the other hand, it improves the model's robustness and adaptability across different data distributions, thereby achieving more balanced recognition of minority classes and a stable overall performance improvement.

5.4. Ablation Experiment

To further explore the impact of IQ signal adaptive rotation, respiratory signal selection, denoising strategy, and augmentation function, we conduct ablation experiments to prove the effectiveness of each processing module. The results are shown in Table 4 and Table 5.

Obviously, **NcSAD *w/o* Rotation** exhibits a noticeable performance degradation (up to 2.27% for the two-classification task and 1.48% for the three-classification task in ACC), indicating that the IQ signal adaptive rotation method effectively reconstructs accurate phase waveforms, thereby leads to a more precise representation of respiratory motion and significantly enhances the overall performance of the model. For the signal selection algorithm,

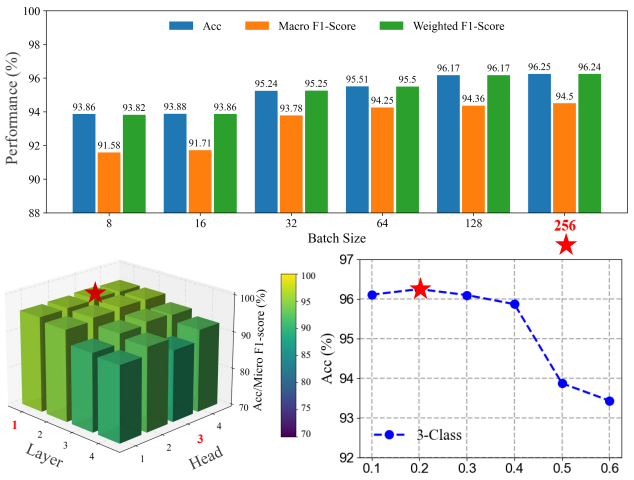


Figure 4. Illustration of hyperparameters and correlation.

when only amplitude or phase signals are used without selective screening, the model's performance drops significantly. This is mainly due to the difference in the ability to perceive amplitude and phase, leading to reduced model stability. In contrast, the signal selection algorithm proposed in this paper can adaptively choose the optimal signal form under current conditions as input, thereby significantly enhancing the model's overall performance.

We also evaluate the effectiveness of the denoising module. Compared with NcSAD, **NcSAD *w/o* Denoising** has significantly performance declined. This is mainly due to that the noise generated by the patient's unconscious movements during sleep can interfere with the actual respiratory waveform, thereby hindering accurate feature extraction and apnea detection. In contrast, the denoising model proposed in this paper effectively filters out outliers and suppresses abnormal signal trends, preserving the underlying respiratory patterns, thereby enables more reliable and robustness apnea detection performance.

Finally, we discuss the effect of the augmentation function. We can observe that for the three-classification tasks, **NcSAD *w* One Aug** has a lower performance than NcSAD (94.02% *V.S.* 96.25% on Acc. In particular, compared to other data modules, the ablation of data augmentation functions leads to a more significant performance degradation. These phenomena indicate that the SSA-based augmentation function can adaptively decompose non-linear signals and achieve multi-scale reconstruction, thereby effectively enhance the model performance.

5.5. Evaluation on Different Hyperparameters

In this section, we investigate the influence of key hyperparameters on three-classification tasks, including the batch size, the number of self-attention heads and layers, and the temperature parameter τ. The corresponding results are illustrated in Fig. 4. It can be observed that performance is

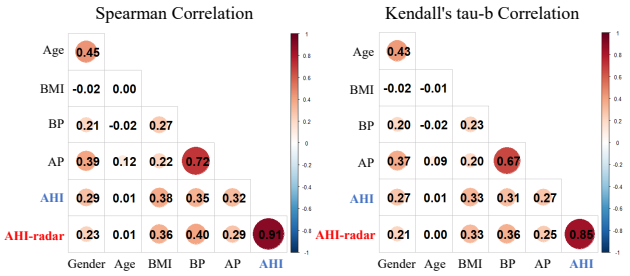


Figure 5. Correlation index analysis.

optimal when the batch size is 256., which is consistent with the finding in [3] that larger batch sizes generally lead to better representation quality in contrastive learning. In addition, the parameters τ, the number of attention heads and layers significantly affect the model performance. Specifically, the model achieves optimal results when τ is set to 0.2, the number of attention heads to 3, and the number of layers to 1, indicating that a moderate network complexity and well-calibrated contrastive temperature jointly contribute to effective feature learning. Notably, the hyperparameter settings used for two-classification tasks are the same as three-classification tasks.

5.6. Correlation between Sleep Apnea and Physiological Factors

In this section, we discuss the correlation between sleep apnea and physiological factors, including gender, age, body mass index (BMI), and blood pressure before and after sleep (BP and AP). Different from the standard Apnea-Hypopnea Inde (AHI) calculation that is based on sleep apnea and hypopnea events from PSG, in this article, we only focus on obstructive sleep apnea, and estimate AHI with OSA, named AHI-radar. As shown in Fig. 5, the correlation coefficients between the AHI and various physiological factors are all below 0.4, indicating weak correlations. The fundamental reason lies in the inherent sample bias of clinically collected sleep apnea datasets, where approximately two-thirds of the subjects are diagnosed with sleep apnea. It is worth noting that, there exists a strong correlation (0.91) between AHI-radar and AHI, indicating that the AHI-radar is highly consistent with the actual AHI. This result provides compelling evidence that the radar-based apnea detection method implemented in NcSAD can achieve accuracy comparable to that of standard clinical measurements.

6. Conclusion

In this paper, we investigate the feasibility of radar-based sleep apnea detection via subtle motion analysis. To this end, we propose a, non-contact sleep apnea detection (NcSAD) method. Specifically, A series of denoising strategies have been proposed to suppress potential noise during clinical data collection. To eliminate the dependence on costly

manual annotations, a Transformer-based self-supervised contrastive learning network is introduced to capture the global temporal dependencies of respiratory signals without reliance on labeled data. Moreover, a dedicated SSA-based augmentation function is designed for respiratory waveforms to enhance sample diversity while preserving intrinsic semantic consistency across augmented samples. We finally release a radar-based OSA dataset to validate NcSAD. Extensive experiments demonstrate that NcSAD achieves accurate and robust detection of sleep apnea events.

7. Acknowledgments

This work is supported by the National Key R&D Program of China (NO.2024YFB3311600), the Natural Science Foundation of China (62272144), the Anhui Provincial Natural Science Foundation (2408085J040), the Major Project of the Anhui Provincial Science and Technology Breakthrough Program (202423k09020001), and the Fundamental Research Funds for the Central Universities (JZ2024AHST0337).

References

[1] Manuel Carro-Domínguez, Stephanie Huwiler, Stella Oberlin, Timona Leandra Oesch, Gabriela Badii, Anita Lüthi, Nicole Wenderoth, Sarah Nadine Meissner, and Caroline Lustenberger. Pupil size reveals arousal level fluctuations in human sleep. *Nature Communications*, 16(1):2070, 2025. 1

[2] Hnin Thiri Chaw, Sinchai Kamolphiwong, and Krongthong Wongsritrang. Sleep apnea detection using deep learning. *Tehnički glasnik*, 13(4):261–266, 2019. 2

[3] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In *International conference on machine learning*, pages 1597–1607. PMLR, 2020. 5, 8

[4] Tian Chen, Jingtao Zhang, Zeju Xu, Stephen J. Redmond, Nigel H. Lovell, Guanzheng Liu, and Changhong Wang. Energy-efficient sleep apnea detection using a hyperdimensional computing framework based on wearable bracelet photoplethysmography. *IEEE Transactions on Biomedical Engineering*, 71(8):2483–2494, 2024. 1, 2

[5] Tianyu Gao, Xingcheng Yao, and Danqi Chen. Simcse: Simple contrastive learning of sentence embeddings. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, EMNLP 2021*, pages 6894–6910, 2021. 5

[6] Xiang Gao, Yanru Li, Wen Xu, and Demin Han. Diagnostic accuracy of level iv portable sleep monitors versus polysomnography for pediatric obstructive sleep apnea: a systematic review and meta-analysis. *Sleep Medicine*, 87: 127–137, 2021. 1

[7] Ronny Gerhard Guendel, Nicolas C Kruse, Francesco Fioranelli, and Alexander Yarovoy. Multipath exploitation for human activity recognition using a radar network. *IEEE Transactions on Geoscience and Remote Sensing*, 62:1–13, 2024. 2

[8] Jie Gui, Tuo Chen, Jing Zhang, Qiong Cao, Zhenan Sun, Hao Luo, and Dacheng Tao. A survey on self-supervised learning: Algorithms, applications, and future trends. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(12):9052–9071, 2024. 5

[9] Dan Guo, Kun Li, Bin Hu, Yan Zhang, and Meng Wang. Benchmarking micro-action recognition: Dataset, methods, and applications. *IEEE Transactions on Circuits and Systems for Video Technology*, 34(7):6238–6252, 2024. 1

[10] Weiqiao Han, Shaozhang Dai, and Mehmet Rasit Yuce. Real-time contactless respiration monitoring from a radar sensor using image processing method. *IEEE Sensors Journal*, 22(19):19020–19029, 2022. 3

[11] Hossein Hassani. Singular spectrum analysis: methodology and comparison. 2007. 5

[12] Shahrokh Javaheri and M Safwan Badr. Central sleep apnea: pathophysiologic classification. *Sleep*, 46(3):zsac113, 2023. 1

[13] Takato Koda, Shigeaki Okumura, Hirofumi Taki, Satoshi Hamada, Hironobu Sunadome, Susumu Sato, Kazuo Chin, and Takuya Sakamoto. Noncontact detection of sleep apnea using radar and expectation-maximization algorithm. *IEEE Sensors Journal*, 2024. 2

[14] Hyun Bin Kwon, Dongyeon Son, Dongseok Lee, Heenam Yoon, Mi Hyun Lee, Yu Jin Lee, Sang Ho Choi, and Kwang Suk Park. Hybrid cnn-lstm network for real-time apnea-hypopnea event detection based on ir-uwb radar. *IEEE Access*, 10:17556–17564, 2021. 2

[15] Shinjae Kwon, Hyeon Seok Kim, Kangkyu Kwon, Hodam Kim, Yun Soung Kim, Sung Hoon Lee, Young-Tae Kwon, Jae-Woong Jeong, Lynn Marie Trotti, Audrey Duarte, et al. At-home wireless sleep monitoring patches for the clinical assessment of sleep quality and sleep apnea. *Science Advances*, 9(21):eadg9671, 2023. 1

[16] Vivianne Landry, Koorosh Semsar-Kazerooni, Tanya Chen, Joshua Gurberg, Lily HP Nguyen, and Evelyn Constantin. Diagnostic accuracy of portable sleep monitors in pediatric sleep apnea: A systematic review. *Sleep Medicine Reviews*, 78:101991, 2024. 1

[17] Jeremy Levy, Daniel Álvarez, Félix Del Campo, and Joachim A Behar. Deep learning for obstructive sleep apnea diagnosis based on single channel oximetry. *Nature Communications*, 14(1):4881, 2023. 1

[18] Kun Li, Dan Guo, Guoliang Chen, Chunxiao Fan, Jingyuan Xu, Zhiliang Wu, Hehe Fan, and Meng Wang. Prototypical calibrating ambiguous samples for micro-action recognition. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 4815–4823, 2025. 1

[19] Siheng Li, Zhi Wang, Fusang Zhang, and Beihong Jin. Fine-grained respiration monitoring during overnight sleep using ir-uwb radar. In *International Conference on Mobile and Ubiquitous Systems: Computing, Networking, and Services*, pages 84–101. Springer, 2021. 3

[20] Siheng Li, Zhi Wang, Beihong Jin, Fusang Zhang, Xiaoyong Ren, and Yitong Zhang. Sleep respiration monitoring using attention-reinforced radar signals. In *2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 1612–1615. IEEE, 2022. 2, 3

[21] Siheng Li, Beihong Jin, Zhi Wang, Fusang Zhang, Xiaoyong Ren, and Haiqin Liu. Leveraging attention-reinforced uwb signals to monitor respiration during sleep. *ACM Transactions on Sensor Networks*, 20(5):1–28, 2024. 1, 2

[22] Siheng Li, Beihong Jin, Fusang Zhang, Zhi Wang, Junqi Ma, Xiaoyong Ren, and Haiqin Liu. Hypnos: A contactless sleep stage monitoring system using uwb signals. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 8(3):1–27, 2024. 3

[23] Xin Liu, Biao Qian, Haipeng Liu, Dan Guo, Yang Wang, and Meng Wang. Seeking false hard negatives for graph contrastive learning. *IEEE Transactions on Circuits and Systems for Video Technology*, 34(8):7454–7466, 2024. 5

[24] Atul Malhotra, Ronald R Grunstein, Ingo Fietze, Terri E Weaver, Susan Redline, Ali Azarbarzin, Scott A Sands, Richard J Schwab, Julia P Dunn, Sujatro Chakladar, et al. Tirzepatide for the treatment of obstructive sleep apnea and obesity. *New England Journal of Medicine*, 391(13):1193–1205, 2024. 1

[25] Paul E Peppard, Terry Young, Mari Palta, and James Skatrud. Prospective study of the association between sleep-disordered breathing and hypertension. *New England Journal of Medicine*, 342(19):1378–1384, 2000. 1

[26] Winfried Randerath, Sébastien Baillieul, and Renaud Tamisier. Central sleep apnoea: not just one phenotype. *European Respiratory Review*, 33(171), 2024. 1

[27] Farjana Snigdha, Shekh MM Islam, Olga Boric-Lubecke, and Victor Lubecke. Obstructive sleep apnea (osa) events classification by effective radar cross section (ercs) method using microwave doppler radar and machine learning classifier. In *2020 IEEE MTT-S International Microwave Biomedical Conference (IMBioC)*, pages 1–3. IEEE, 2020. 2

[28] Syed Doha Uddin, Md Shafkat Hossain, Shekh MM Islam, and Victor Lubecke. Heart rate variability-based obstructive sleep apnea events classification using microwave doppler radar. *IEEE Journal of Electromagnetics, RF and Microwaves in Medicine and Biology*, 7(4):416–424, 2023. 2, 6

[29] A Vaswani. Attention is all you need. *Advances in Neural Information Processing Systems*, 2017. 5

[30] Jia Wang, Jiangtao Xue, Yang Zou, Yuxin Ma, Junhan Xu, Yanming Li, Fei Deng, Yiqian Wang, Kai Xing, Zhou Li, et al. A dual-modal wearable pulse detection system integrated with deep learning for high-accuracy and low-power sleep apnea monitoring. *Advanced Science*, page 2501750, 2025. 1

[31] Tongzhou Wang and Phillip Isola. Understanding contrastive representation learning through alignment and uniformity on the hypersphere. In *International Conference on Machine Learning*, pages 9929–9939. PMLR, 2020. 5

[32] Xinke Wang, Jingyuan Xu, Xiao Sun, Mingzheng Li, Bin Hu, Wei Qian, Dan Guo, and Meng Wang. Facial depression estimation via multi-cue contrastive learning. *IEEE Transactions on Circuits and Systems for Video Technology*, 35 (6):6007–6020, 2025. 5

[33] Zhi Wang, Beihong Jin, Siheng Li, Fusang Zhang, and Wenbo Zhang. Ecg-grained cardiac monitoring using uwb signals. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 6(4):1–25, 2023. 3

[34] Linshan Wu, Jiaxin Zhuang, and Hao Chen. Voco: A simple-yet-effective volume contrastive learning framework for 3d medical image analysis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 22873–22882, 2024. 5

[35] Yingxiao Wu, Haocheng Ni, Changlin Mao, Jianping Han, and Wenyao Xu. Non-intrusive human vital sign detection using mmwave sensing technologies: A review. *ACM Transactions on Sensor Networks*, 20(1):1–36, 2023. 2

[36] Baile Xie and Hlaing Minn. Real-time sleep apnea detection by classifier combination. *IEEE Transactions on Information Technology in Biomedicine*, 16(3):469–477, 2012. 2

[37] Shufan Yang, Julien Le Kernec, Olivier Romain, Francesco Fioranelli, Pierre Cadart, Jérémy Fix, Chenfang Ren, Giovanni Manfredi, Thierry Letertre, Israel David Hinostroza Sáenz, et al. The human activity radar challenge: benchmarking based on the 'radar signatures of human activities' dataset from glasgow university. *IEEE Journal of Biomedical and Health Informatics*, 27(4):1813–1824, 2023. 2

[38] Yuyan Zhang, Zhe Liu, Chunjie Yang, Xiaoke Huang, Siwei Lou, Hanwen Zhang, and Duojin Yan. Unveiling dynamics changes: Singular spectrum analysis-based method for detecting concept drift in industrial data streams. *Knowledge-Based Systems*, 293:111640, 2024. 5

[39] Tianyue Zheng, Zhe Chen, Shujie Zhang, and Jun Luo. Catch your breath: Simultaneous rf tracking and respiration monitoring with radar pairs. *IEEE Transactions on Mobile Computing*, 22(11):6283–6296, 2022. 3