Expert Systems With Applications 289 (2025) 128360

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Noncontact capacitive coupling ECG-Derived respiratory signals using the conformer based time–frequency domain generative adversarial network

Zhijun Xiao [a,b], Maarten De Vos [b,c], Christos Chatzichristos [b], Kejun Dong [e], Yunyi Jiang [a], Zhongyu Wang [a], Yuwei Zhang [f], Fei Ding [d], Chenxi Yang [a], Jianqing Li [a], Chengyu Liu [a,*]

[a] The State Key Laboratory of Bioelectronics, School of Instrument Science and Engineering, Southeast University, Nanjing China
[b] The Department of Electrical Engineering, KU Leuven, Belgium
[c] The Department of Development and Regeneration, KU Leuven, Belgium
[d] The college of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing, China
[e] Center for Data Science, Nell Hodgson Woodruff School of Nursing, Emory University, Atlanta, GA, United States
[f] the School of Computer Science, Jiangsu University of Science and Technology, Zhenjiang, China

ARTICLE INFO

Keywords:
Bedside Respiratory Monitoring
Capacitive coupling electrocardiogram
Generative Adversarial Network
Time-Frequency Domain Enhancement

ABSTRACT

Respiratory monitoring and analysis is a key method for detecting sleep-related diseases. This paper presents a novel approach for respiratory monitoring that utilizes noncontact capacitive coupling electrocardiograms-derived respiration (cEDR) method. We propose a Time-Frequency Domain Generative Adversarial Network (TF-GAN) method for generating respiratory signals, and successfully apply it to capacitive coupling electrocardiograms(cECG). First, we analyze the mechanism of respiratory coupling with cECG and verify the feasibility of the theory. Then, using the developed device, we collect cECG data from 16 subjects during the night and simultaneously collect respiratory signals as a reference, to validate the feasibility of our approach. Next, we convert the collected cECG data into time–frequency domain features using Short-Time Fourier Transform (STFT) and input these features into a Convolution-augmented transformer (Conformer) based Generative Adversarial Network(GAN) to generate the cEDR. The network architecture integrates self-attention mechanisms and time–frequency domain enhancement mechanisms to effectively extract the respiratory energy components. Finally, we compare the generated respiratory signals with the reference signals. The experimental results show that the generated respiratory signals exhibit a high correlation with the reference signals. Specifically, 86.3 % of the signals have a absolute waveform correlation coefficient greater than 0.5, indicating good reproduction of real breathing waveforms. Our proposed model demonstrates superior performance in respiratory signal extraction, achieving a low Root Mean Square Error (RMSE) of 0.96 ± 0.12 bpm and a high agreement rate of 94.83 % ± 0.30 % within the Bland–Altman limits. Additionally, the model maintains an effective respiratory segment ratio of 67.56 % ± 8.89 %, even under poor cECG signal conditions, showcasing its robustness and reliability.

1. Introduction

Respiratory monitoring during sleep is essential for health assessment, as it can help detect potential health issues such as sleep apnea, which is also recognized as being closely associated with cardiovascular diseases (Dong et al., 2022). Accurate monitoring of respiratory activity aids in the early detection of diseases, guides treatment plans, and improves overall sleep quality. Noncontact monitoring technologies offer patients a more comfortable, interference-free approach, especially suitable for long-term home-based health monitoring (Jin et al., Nov 2017). Noncontact monitoring technologies, such as radar-based (W. Do, R. Russell, C. Wheeler, M. Lockwood, M. De Vos, I. Pavord, and M. Bafadhel, "Performance of Contactless Respiratory Rate Monitoring by Albus Home(TM), an Automated System for Nocturnal Monitoring at Home: A Validation Study," Sensors (Basel), vol. 22, no. 19, Sep 21, 2022) and piezoelectric sensors (Chiu et al., 2013), offer patients a more comfortable, interference-free approach, making them particularly suitable for long-term home-based respiratory monitoring.

Common techniques for respiratory monitoring during sleep include airflow sensors (Jin et al., 2022), impedance pneumography (Song et al.,

* Corresponding author.
E-mail address: chengyu@seu.edu.cn (C. Liu).

https://doi.org/10.1016/j.eswa.2025.128360
Received 5 March 2025; Received in revised form 16 May 2025; Accepted 26 May 2025
Available online 28 May 2025
0957-4174/© 2025 Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

Aug 2024), and inductance plethysmography sensors (Garbino et al., Feb 2014). However, these methods are often too cumbersome for applications that require comfortable, long-term respiratory monitoring. In addition to the methods mentioned above, extracting respiratory waveforms from other physiological signals is considered a more cost-effective approach and better suited for obtaining respiratory information in wearable monitoring applications (Dong et al., Aug 2021). This has led to the development of a novel respiratory signal generation technique known as electrocardiogram-derived respiration (EDR) (Thomas et al., Sep 2005). EDR enables synchronized respiratory signals to be calculated from electrocardiogram (ECG) data through algorithmic processing. The primary advantage of extracting respiratory information from ECG signals is the ability to monitor both cardiac and respiratory activity simultaneously, allowing for synchronized heart and respiratory monitoring. This approach significantly reduces the number of sensors and wiring required, simplifying device design and wearability, and greatly enhancing user comfort.

Conventional ECG monitoring uses electrodes adhered to the skin, which can accurately record ECG signals. But it is uncomfortable and may cause skin irritation (Uguz et al., 2022). To address this issue, noncontact capacitive coupling ECG (cECG) monitoring technology has emerged. It uses high-impedance circuit sensing to capture ECG signals without direct skin contact, allowing signal acquisition through a layer of fabric. This approach not only enhances user comfort but also avoids skin irritation, making it a promising method for home-based ECG monitoring (Wang et al., Mar 2023; Leicht et al., Apr 2017). Since cECG is essentially an electrophysiological signal, it retains characteristics similar to conventional ECG, enabling comparable monitoring and analysis (Wang et al., 2020).

The theory of extracting respiratory signals from conventional ECG has been validated (Orphanidou et al., Jan 2013). Respiratory modulation of the ECG can occur through several mechanisms, including baseline wander (BW), amplitude modulation (AM), and frequency modulation (FM) (Bailón et al., 2006). Respiratory activity induces changes in thoracic volume, which in turn affects the relative position of the heart's electrical axis, leading to alterations in the ECG waveform, particularly in the amplitude of the R wave and the morphology of the QRS complex. BW is primarily caused by the mechanical coupling between respiratory movements and thoracic impedance. As the chest expands and contracts during breathing, changes in thoracic impedance induce a low-frequency oscillation in the ECG baseline, which is synchronized with the respiratory cycle.

Additionally, during inhalation, changes in intrathoracic pressure cause the atrioventricular node to stretch. As intrathoracic pressure decreases, the left ventricular stroke volume reduces, and heart rate increases. During exhalation, increased vagal activity slows the heart rate. These mechanisms collectively result in significant respiratory modulation of the ECG signal, providing a basis for extracting respiratory information from the ECG (Helfenbein et al., 2014).

The use of ECG waveforms for respiratory disease monitoring has garnered widespread attention, primarily focusing on two directions: directly evaluating respiratory disorders based on ECG waveforms, and extracting respiratory waveforms from ECG signals for analysis. For the first direction, deep learning techniques, have been widely applied. These networks are capable of automatically extracting key features from ECG signals and classifying respiratory disorders. For instance, Xie designed an RNN-based deep network to identify sleep-disordered breathing (Xie et al., Jul 2023); Tyagi utilized an enhanced deep belief network model to identify sleep apnea (Tyagi and Agrawal, Feb 2023). Hu proposed a personalized adjustment method that combines a convolutional neural network-based autoencoder with a semi-supervised algorithm. This approach aims to improve the accuracy of single-lead ECG in the automated detection of obstructive sleep apnea through low-cost personalized fine-tuning (Hu et al., 2023). Directly using ECG waveforms may overlook features that are more sensitive to respiratory activity. Given that respiratory signals have lower frequencies and amplitudes, directly inputting ECG waveforms for the detection of respiratory disorders may result in the loss of critical respiratory information. For the second direction, by extracting specific respiratory waveforms from the ECG signal, clearer features related to respiratory frequency and periodic changes can be obtained, which improves the accuracy of detecting respiratory disorders. The extracted respiratory waveforms not only provide a clearer reflection of the respiratory state but also offer additional physiological parameters, such as breathing rate (BR), and variability analysis. This approach offers advantages for long-term monitoring and refined diagnosis of respiratory issues.

Currently, research on extracting respiratory waveforms using deep learning methods is relatively scarce, with most studies focusing on traditional signal processing techniques. These methods include filtering-based approaches and feature-based approaches. Among the filtering methods, Lindberg use a bandpass filter to extract respiratory signals from ECG (Lindberg et al., 1992). Yi utilized discrete wavelet transform (DWT) to decompose the ECG signal, and then reconstructed the respiratory signal using the ninth decomposition level, which corresponds to the frequency range of 0.2–0.4 Hz (Yi et al.n.d); For feature-based methods, various characteristics of the ECG signal have been extracted to reconstruct respiratory signals, such as the area of the QRS complex (Sadr and de Chazal, Feb 2019); RR intervals, R-wave amplitude (Dong et al., Aug 2021); and the slope of the QRS complex (Lázaro et al., 2014).

For cECG signals, Similarly to the ECG, we can categorize the respiratory influences on cECG into two aspects. First, similar to conventional ECG monitoring, respiratory activity leads to changes in thoracic volume, which, in turn, causes variations in the cECG waveform. Second, the mechanism of capacitive coupling enhances the respiratory coupling. During the cECG recording process, as a layer of clothing is present, the relative displacement between the skin and the electrode due to respiration can affect the impedance at the skin-electrode interface. This variation in impedance can influence the noncontact ECG signal, resulting in baseline drift that is synchronized with the respiratory signal (Uguz et al., 2020). Although cECG signals also exhibit significant respiratory coupling, there has been limited research on cECG-derived respiration (cEDR). we will use inspiration from ECG for adapting it to the specifical properties of the cECG.

To validate the effectiveness cEDR method, this paper constructs a noncontact ECG respiratory database and proposes a Time-Frequency Domain Generative Adversarial Network Network (TF-GAN) for respiratory signal generation. The approach involves two steps: first, converting the signal into time–frequency domain features using Short-Time Fourier Transform (STFT) and then inputting these features into a conformer-based GAN network. The network generates time–frequency feature maps of the respiratory signal, which are then transformed back into the final respiratory signal through the inverse STFT (ISTFT). The network architecture integrates self-attention mechanisms and time–frequency domain enhancement techniques, which consider both amplitude and phase characteristics of the signal, ensuring high-quality generation of respiratory signals.

The main contributions of this paper are:

(1) A theoretical analysis of the coupling mechanism between cECG and respiratory signals is provided, and a cECG and piezoelectric-film-based mattress system for ECG and respiratory monitoring is developed to support subsequent validation. Based on TF-GAN, this paper extracts deep features from both the time and frequency domains to generate respiratory signals. The method not only focuses on the R-wave features in the respiratory signal but also emphasizes baseline features, fully leveraging the amplitude and phase information of the signal.

(2) As far as we know, this paper is the first to use cECG signals to extract respiratory waveforms. Compared to conventional sensors, this method avoids the discomfort, offering a more comfortable, convenient, and sustainable solution for long-term respiratory monitoring during sleep. In addition, cECG signals enable the high-precision monitoring of RR intervals, allowing for synchronized monitoring of both cardiac and

respiratory activities.

(3) Our work emphasizes that during actual signal acquisition, some uncontrollable factors (such as relative displacement between the skin and electrode, electrode position, etc.) may cause the respiratory waveform to exhibit positive or negative correlations. However, despite these interferences, the beat-to-beat accuracy of the respiratory signal remains unaffected, ensuring the reliability and precision of the respiratory signal extraction.

The paper is organized as follows: Section II discusses the signal characteristics of cECG, the coupling mechanisms of respiratory signals, and the structure of the noncontact ECG hardware system used. Section III presents the model design of the TF-GAN Network. Section IV provides the results of the model, comparing the generated respiratory signals with real respiratory signal features. Section V includes a discussion, while Section VI concludes the paper.

2. Principle of cECG respiratory coupling and hardware Implementation

2.1. Skin-electrode interface of cECG

Fig. 1 (a) illustrates the electrode arrangement for cECG, the skin-electrode interface, and the analog frontend (AFE) circuit of cECG used in this study. The strip-shaped fabric electrodes (electrode A and B) are placed below the shoulder and chest area. The skin-electrode interface impedance, Z_s, consists of two components: the first is the skin impedance, Z_{skin}, (including R_0, C_1, *and* R_1). The second component is the coupling impedance, $Z_{coupling}$, which arises due to the parallel combination of the coupling capacitance, C_s, and the leakage resistance, R_{leak}. For Z_{skin}, the impedance is relatively small at the level of serval MΩ. For $Z_{coupling}$, it is usually large (ranging from tens of MΩ to a few GΩ, depending on the thickness, humidity, and some other characteristics of the fabric). So, the skin-electrode coupling impedance, Z_s, can be approximately equal to the impedance generated by the coupling capacitance. In the ideal case, the leakage resistance is neglected. Z_s can be defined as:

$$Z_s = Z_{skin} + Z_{coupling} \approx \frac{1}{SC_s} \approx \frac{d + \Delta d(t)}{S\varepsilon_r\varepsilon_0 A_s} \tag{1}$$

where C_s is the coupling capacitance; A_S is the area of the skin-electrode contact area; d is the thickness of the fabric, $\Delta d(t)$ is the relative displacement between the skin and the electrode mainly caused by respiratory movement; ε_0 is the vacuum dielectric constant, ε_r is the relative dielectric constant of clothing.

2.2. Baseline changes in cECG induced by respiration

For electrode A, as we can see in Fig. 1(a), when v_a passes through the skin-electrode interface, the input voltage $v_{a'}$ to the buffer can be expressed as:

$$v_{a'} = \frac{Z_{in}}{Z_{in} + Z_s} v_a = \frac{Z_{in}}{Z_{in} + \frac{d + \Delta d_a(t)}{S\varepsilon_r\varepsilon_0 A_s}} v_a \tag{2}$$

this is the same as electrode B. So, v_{cECG} can be defined as:

$$v_{cECG} = v_{a'} - v_{b'} = \frac{Z_{in}}{Z_{in} + \frac{d + \Delta d_a(t)}{S\varepsilon_r\varepsilon_0 A_s}} v_a - \frac{Z_{in}}{Z_{in} + \frac{d + \Delta d_b(t)}{S\varepsilon_r\varepsilon_0 A_s}} v_b \tag{3}$$

where, v_a and v_b are the potentials at the contact points of electrodes A and B on the body surface. $v_{a'}$ and $v_{b'}$ is the input potential of the buffer. Z_{in} is the impedance of AFE. $\Delta d_a(t)$ and $\Delta d_b(t)$ are the relative displacement between the skin and the electrode mainly caused by respiratory movement.

During inhalation and exhalation stage, the relative displacement of the skin electrodes affects the impedance of the skin-electrode interface, thereby modulating the cECG signal. This modulation is primarily reflected in the overall potential shift of the noncontact cECG, manifesting as a significant baseline drift in the waveform morphology.

2.3. Morphological changes in cECG waveform induced by respiration

During respiration, changes in chest cavity volume affect the position

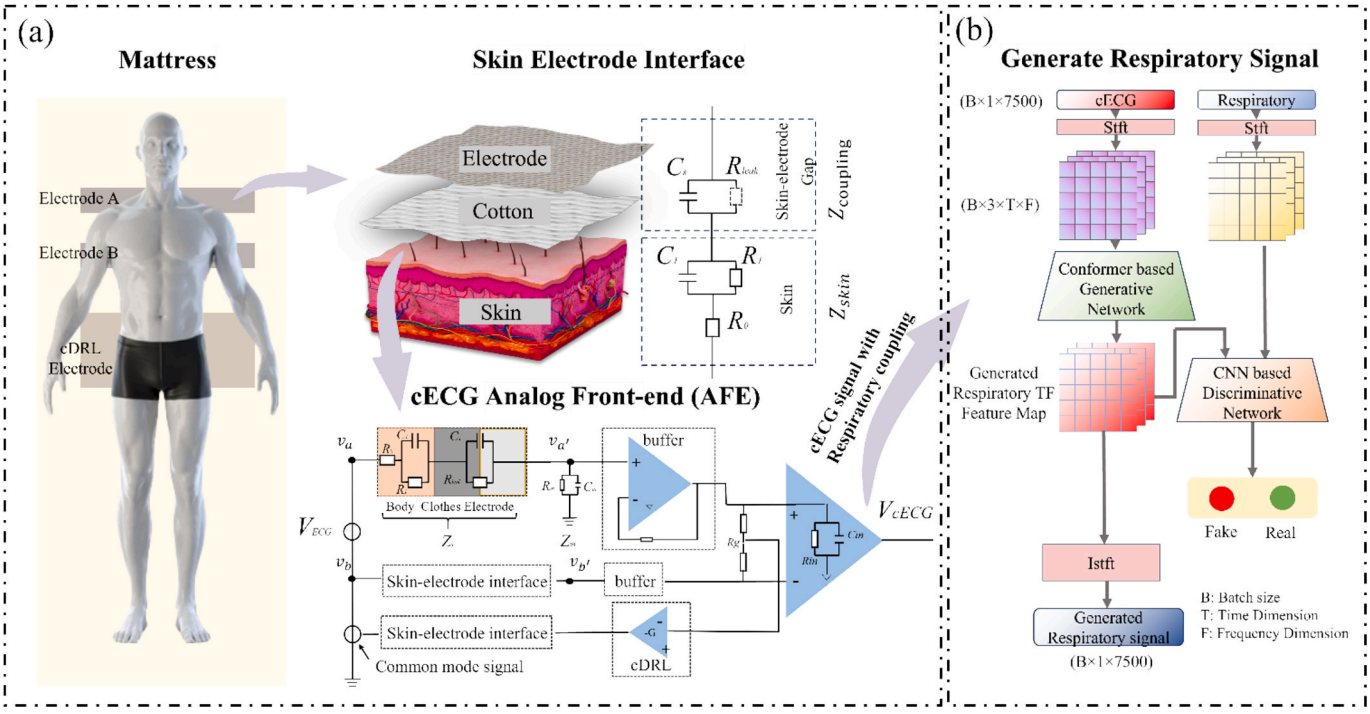


Fig. 1. The schematic diagram of the noncontact cECG acquisition principle and the structure of the proposed GAN algorithm, (a) illustrates the electrode placement for noncontact cECG, the skin-electrode interface, and the analog frontend circuit of cECG; (b) depicts the architecture of the proposed GAN network.

of the heart's electrical axis, thereby altering the ECG waveform, especially the amplitude of the R-wave and the morphology of the QRS complex. Additionally, during inhalation, changes in thoracic pressure stretch the atrioventricular node, leading to a reduction in left ventricular blood ejection volume and an increase in heart rate; whereas during exhalation, increased vagal nerve activity slows the heart rate. These mechanisms together contribute to the rich respiratory modulation present in the ECG signal, for example, amplitude of the R-wave (Fleming and Tarassenko, 2007); the RR interval (Shelley et al., 2006), and the slope of the QRS complex wave (Garde et al., 2014). In the process of cECG signal acquisition, these changes already exist before the ECG signal reaches the skin-electrode interface. So, such kind of respiratory modulation is the same as cECG signal.

Therefore, the modulation of cECG signals by respiration consists of two main components: the first is the modulation of the ECG signal due to changes in the relative position of the heart's electrical axis. The second is the modulation of the cECG signal caused by the relative movement of the skin-electrode interface (capacitive coupling interface) when the ECG signal is transmitted to the skin-electrode interface.

2.4. cECG system and respiratory sensor used in This work

To record respiratory and cECG signals during sleep, a noncontact cECG and respiratory signal joint acquisition system was established. This system allows for the synchronous acquisition of both cECG and respiratory signals. The system and experimental setup used are shown in Fig. 2(a). The details of the two modules are as follows:Fig. 3..

2.4.1. cECG module

The cECG system used in this study consists of a single-gain high-impedance buffer module, a differential amplification circuit module, and an AD converter. The single-gain high-impedance buffer module is composed of a voltage follower built with LTC6268, known for its extremely high input impedance and low input bias current, facilitating the recording of bioelectric potentials. The differential amplification circuit module performs the differential operation on the bioelectric potentials sensed by two electrodes, ultimately yielding the cECG signal. The AD converter converts the analog signal into digital data, facilitating subsequent digital recording and storage. For specific hardware technical details, please refer to our previous work (Xiao et al., 2023).

2.4.2. Respiratory module

The respiratory signal recorded module is composed of a polyvinylidene fluoride (PVDF) sensor and a high input impedance charge amplifier. The charge amplifier records and amplify the respiratory signals by sensing the charge generated by the PVDF sensor. The charge amplifier INA332 (Texas Instruments Inc., USA) uses a single power supply. High-precision gain amplification is achieved by configuring peripheral resistors and capacitors.

3. Materials and methodology

In this study, a noncontact cECG respiratory database was created by collecting night-time noncontact cECG and reference respiratory signals from 16 subjects. The cECG data were input into a TF-GAN network to generate respiratory signals. The generated signals were then compared with reference respiratory signals, confirming the method's effectiveness. The network's structure is shown in Fig. 1(b).

3.1. Experimental Preparation

3.1.1. cEDR dataset

Due to the lack of existing cECG and respiratory signal datasets, we created a new dataset. In the dataset, 16 subjects were tested while wearing cotton T-shirts and lying freely on a designed noncontact cECG smart mattress. Both cECG and respiratory signals were synchronously recorded with a sampling rate of 500 Hz. Specific information and collection duration for each subject are shown in Table 1. The experimental protocols were approved by the Ethics Committee of the First Affiliated Hospital of Nanjing Medical University (Reference code: 2020-SRFA-183).

3.1.2. Signal Denoising and Segmentation

To reduce the impact of low-frequency and high-frequency vibration

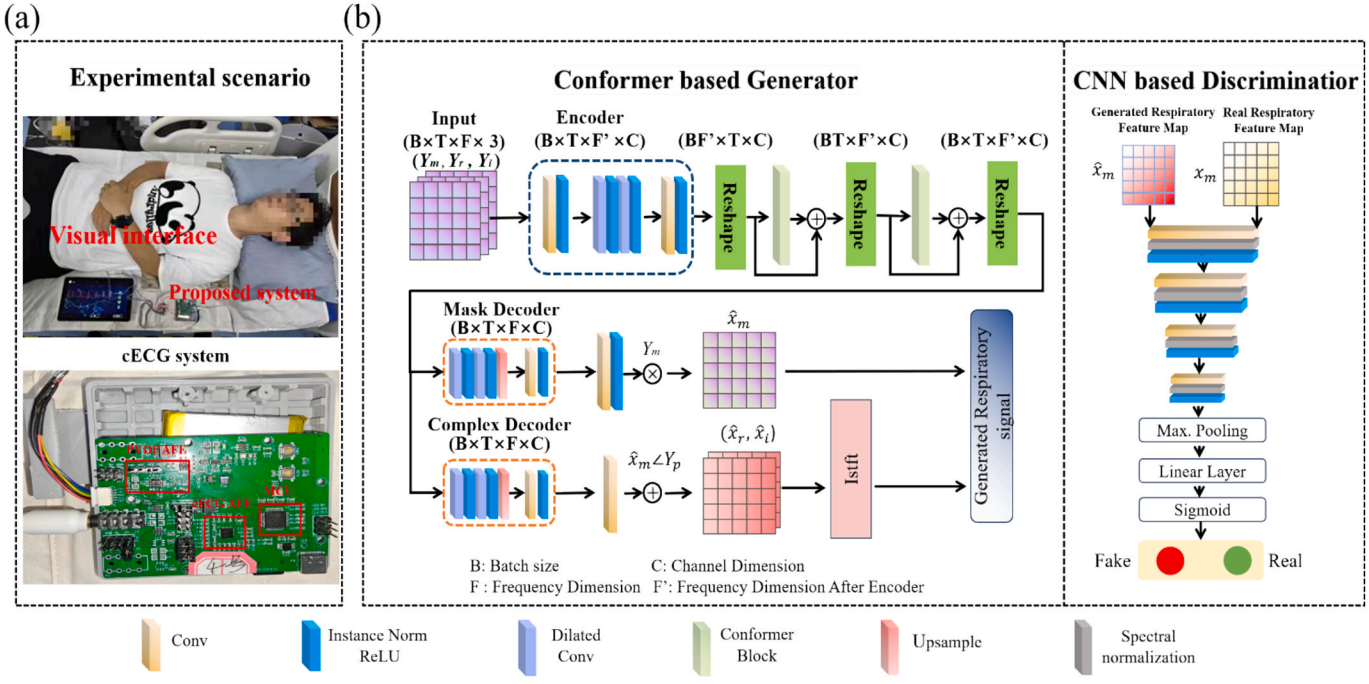


Fig. 2. The experimental setup and the algorithm schematic used in the study. (a) the experimental setup diagram and the cecg hardware system; (b) the network structure of the generator and discriminator in the proposedGenerative Adversarial Network (GAN).

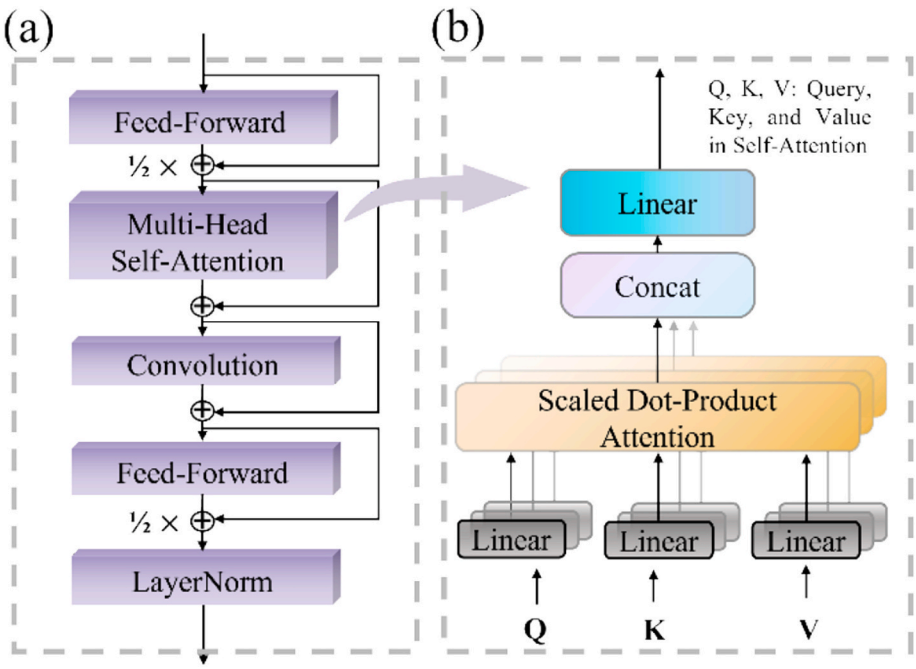


Fig. 3. Architecture of conformer block. (a) is the conformer block; (b) is the Multi-Head Self-Attention Block.

Table 1
Dataset for Respiratory Accuracy Evaluation (Simultaneous Recording of cECG and Respiratory Signals).

Subjects	Gender	Recording Time [h]	(Age [years], BMI [kg/m^2])	Subjects	Gender	Recording Time [h]	(Age [years], BMI [kg/m^2])
1	male	7.07	(30, 25.0)	9	male	7.52	(34, 23.3)
2	male	4.01	(25, 22.4)	10	male	6.13	(42, 17.5)
3	male	7.30	(27, 20.5)	11	male	6.34	(53, 22.8)
4	male	7.14	(34, 27.6)	12	male	8.79	(18, 17.2)
5	female	3.95	(27, 24.6)	13	female	8.23	(55, 26.0)
6	female	7.88	(26, 20.3)	14	female	7.23	(31, 17.7)
7	female	5.90	(35, 18.2)	15	female	5.84	(40, 18.0)
8	female	7.62	(25, 17.7)	16	female	7.62	(48, 25.4)

noise on the respiratory signal, the respiratory signal is bandpass filtered in the range of 0.15–0.4 Hz, which corresponds to the normal breathing rate range of 9–24 breaths per minute (bpm) (Quintana et al., 2016). Similarly, since cECG signals are also affected by baseline drift, power line interference (PLI), and high-frequency noise, we applied a 6th-order Butterworth bandpass filter with a passband of 0.15–35 Hz. To further reduce computational complexity, the original signals were downsampled to a sampling rate of 250 Hz. Following these preprocessing steps, both the respiratory and cECG signals were segmented into 30-second segments with a 10-second overlapping window.

3.2. Overview of the proposed model

GANs are deep learning models consisting of a generator and a discriminator. The unique structure of GANs makes them highly effective in generating realistic data, and they have been widely used in cross-modal generation tasks. In this paper, we propose a TF-GAN network. The overall algorithm flow is shown in Fig. 2(b). The key components of our architecture, including: STFT processing, encoder modules, conformer blocks, mask decoder, complex decoder, and the discriminator design is shown in Table 2. The processing flow of the generator in the GAN network includes the following steps: first, a STFT is applied to the cECG signal segment to obtain the complex spectrogram feature map, Y_{input}. Then, an encoder based on DenseNet is used to encode Y_{input}. Next, the encoded feature map undergoes time–frequency feature enhancement based on the conformer module. Finally, the signal is passed through a decoder and an ISTFT to generate the respiratory signal. The discriminator in TF-GAN architecture is primarily based on a convolutional network, which is designed to distinguish real signals from generated ones.

This study employs GANs for respiratory signal generation, leveraging their unique adversarial training mechanism to effectively learn the complex cross-modal mapping from cECG to respiratory signals. The time–frequency domain processing approach (STFT) was selected due to its ability to effectively decouple respiratory frequency components from ECG signals, significantly reducing model complexity. The time–frequency analysis captures not only temporal features but also frequency-domain characteristics, which substantially facilitates subsequent respiratory signal extraction.

3.2.1. Short-Time Fourier Transform (STFT)

For a cECG signal segment $y \in R^{L\times 1}$, the STFT transforms the signal into a complex-valued spectrogram $Y \in R^{T\times F\times 2}$, where T and F represent the time and frequency dimensions, respectively. The spectrogram can be expressed as (Abdulatif et al., 2024):

$$Y = Y_m e^{jY_p} = Y_r + jY_i \tag{4}$$

where Ym, Yp, Yr and Yi represent the magnitude, phase, real, and imaginary components of the spectrogram, respectively. To form the input for the generator, the real and imaginary components (Yr and Yi) are combined with the magnitude (Ym) by concatenation. Thus, the input for a single sample can be written as:

Table 2
Parameter Configurations of Key Modules in Generator and Discriminator.

Module	**Main Submodules**	Parameter
Generator	STFT	Window length = 256, stride = 60.
	Encoder	Conv1: kernel= (1,1), stride= (1,1), output_channels = 64 DilatedDenseNet: depth = 2, input_channels = 64 Conv2: kernel= (1,3), stride= (1,2), padding= (0, 1), output_channels = 64
	Conformer	Attention: (1) dim = 64, dim_head = 16, heads = 4, dropout = 0.2, Conv block: (1) Conv3: output_channels = 256, kernel = 1 stride = 1; (2) DepthwiseConv1d: output_channels = 128, kernel = 31 stride = 1; (3) Conv4: output_channels = 64, kernel = 1 stride = 1.
	Mask Decoder and Conv	DilatedDenseNet: depth = 2, input_channels = 64; Conv5: kernel= (1, 2), stride= (1, 1), output_channels = 1
	Complex_Decoder + Conv	DilatedDenseNet: depth = 2, input_channels = 64; Conv6: kernel= (1, 2), stride= (1, 1), output_channels = 1
Discriminator	Conv block in Discriminator	Conv7: kernel= (4,4), stride= (2,2), output_channels = 16 Conv8: kernel= (4,4), stride= (2,2), output_channels = 32 Conv9: kernel= (4,4), stride= (2,2), output_channels = 64 Conv10: kernel= (4,4), stride= (2,2), output_channels = 128

$$Y_{input} = concat(Y_r, Y_i, Y_m) \tag{5}$$

Concatenating amplitude information (Y_m) with real (Y_r) and imaginary (Y_i) components as input accelerates the learning process, helping the network better capture the relationship between amplitude and phase, and avoiding unnecessary complexity in computation.

3.2.2. Generator architecture

3.2.2.1. Encoder and decoder. Encoder-decoder is a type of artificial neural network architecture used to learn efficient embeddings of data. The primary function of the encoder is to extract features from the input cECG signal, which is first transformed into a spectral representation. This process may involve multiple convolutional layers, pooling layers, and fully connected layers, which progressively extract higher-level features from the input signal. The primary function of the decoder is to convert the encoded features, generated by the encoder and enhanced through intermediate layers, back into the target output signal—specifically, the respiratory signal's spectral representation. The encoder and decoder used in this paper are based on Dilated DenseNet (Pandey et al. n.d).

3.2.2.2. Time and frequency encoder. The input of the feature is $Y_{input} \in R^{B\times T\times F\times 3}$, B is the batch size. The encoder is a network structure that integrates Convolution-DenseNet-Convolution layers. Each convolutional block consists of a convolutional layer, followed by instance normalization and a ReLU activation function. The initial convolutional layer expands the number of input channels from 3 to 64. The dilated DenseNet architecture incorporates four convolutional blocks interconnected through dense connections. Dilated convolutions effectively expand the receptive field while maintaining the same number of kernels and layers. The final convolution layer is used to reduce the number of output channels from the dilated DenseNet, thus decreasing the model's complexity.

3.2.2.3. Conformer block. The Conformer block is a hybrid model architecture consisting of four components in sequence: a feed-forward layer, a multi-head self-attention (MHSA) layer, a convolutional layer, and a second feed-forward layer followed by a LayerNorm layer. The conformer architecture combines the Transformer specialized in extracting global features and the convolutional layer specialized in exploring local features, would be well-suited for capturing evidence of respiratory form cECG time and frequency feature map.

First, for the feed-forward layer, two linear transformation layers and an activation function are used to introduce non-linearity, and residual connections and layer normalization are applied to improve training efficiency and stability. This module is particularly important in the Conformer architecture, as it provides the model with the ability to handle complex data. Inspired by Macaron-Net, this work employs half-step residual weights in the feed-forward (FFN) modules (Chan and Chin, 2021). This half-step FFN maintains numerical stability during the computation process and controls the balance and stability of model training.

For the MHSA layer, which is widely used in Transformer architectures, we integrate relative sinusoidal position encoding and use pre-normalized residual units along with dropout to enhance the robustness of the model.

The convolutional layer is a sequential module (Gulati et al., 2005) that first applies layer normalization to the input 64-dimensional channel data, then rearranges the channel dimension to the middle position through axis permutation. Next, a 1x1 convolution expands the channels to 256 dimensions, followed by compression back to 128 dimensions via GLU gating mechanism. Subsequently, depthwise separable convolution with kernel size 31 extracts sequence features, which then undergoes batch normalization and Swish activation to enhance nonlinearity (Ramachandran et al., 1710). Finally, a 1x1 convolution reduces channels back to 64 dimensions while restoring the original dimension order. The overall design efficiently integrates channel and sequence information through convolutional and gating mechanisms, making it suitable for temporal feature extraction. Mathematically, this means, for input p to a Conformer block, the output y_{out} of the block is:

$$\begin{aligned} \widetilde{p} &= p_i + \frac{1}{2}FFN(p) \\ p' &= \widetilde{x}_i + MHSA(\widetilde{p}) \\ p'' &= p'_i + Conv(p') \\ y_{out} &= Layernorm\left(p'' + \frac{1}{2}\mathrm{FFN}(p'')\right) \end{aligned} \tag{6}$$

where FFN refers to the Feed forward module, MHSA refers to the Multi-Head Self-Attention module, and Conv refers to the Convolution module as described in the preceding sections.

3.2.2.4. Time and frequency domain enhancement based on conformer block. We use two Conformer modules sequentially: the first stage captures time dependencies, and the second stage captures frequency dependencies. First, we adjust the input feature $D_{feat} \in R^{B\times T\times F'\times C}$ to $D_{feat}{}^{T} = R^{BF'\times T\times C}$ to extract time-domain features. The output is then element-wise added with $D_{feat}{}^{T}$ to obtain $D_{feat_o}{}^{T}$, and the dimensions are adjusted back to $D_{feat}{}^{F} = R^{BT\times F'\times C}$. After passing through the Conformer blocks with residual connections, we obtain the final output, which is enhanced in both time and frequency domains.

3.2.2.5. Time and frequency decoder and respiratory signal Reconstruction. As is shown in Fig. 2(b), The input to the decoder network is the feature map enhanced by the conformer block. The decoder consists of two parts: the first is the magnitude decoder, and the

second is the complex decoder, aimed at improving the model's ability to handle complex spectrograms. Both parts pass through a dilated DenseNet, with the same structure as the encoder, and then perform upsampling on the frequency dimension, followed by convolution, instance normalization, and ReLU activation.

For the magnitude decoder block, a convolutional layer is used to reduce the number of channels to 1, followed by another convolutional layer with a ReLU activation function to predict the final magnitude.

For the complex decoder, in order to incorporate the phase information of the original signal, following the approach in reference (Li et al., 2021); $\hat{x}_m$ is merged with the cECG signal, Y_p, to produce the magnitude-enhanced complex spectrogram. Then, this enhanced complex spectrogram is element-wise added to the output of the complex decoder $(\hat{x}_r^{'}, \hat{x}_i^{'})$. Thus, the final complex spectrogram can be defined as:

$$\hat{X}_r = \hat{X}_m \cos(Y_p) + \hat{X}_r^{'} \tag{7}$$

$$\hat{X}_i = \hat{X}_m \sin(Y_p) + \hat{X}_i^{'} \tag{8}$$

where, $\hat{X}_r$ and $\hat{X}_i$ represent the real and imaginary parts of the final generated respiratory signal's complex spectrogram. We can use an ISTFT to get the finally respiratory signal $\hat{X}$.

3.3. Discriminator architecture

The discriminator in our GAN framework is meticulously designed to handle cECG data and produce precise respiratory signals. It consists of four convolutional blocks, each starting with a convolution layer, followed by instance normalization and ReLU activation. After passing through the convolutional blocks, a global average pooling layer is utilized to reduce the spatial dimensions, ensuring computational efficiency. Two feed-forward layers then refine the extracted features, and a sigmoid activation function is applied at the final stage to generate the output. This design improves the discriminator's ability to accurately differentiate between real and generated respiratory signals.

3.4. Loss function

This work employs a multi-loss calculation approach, where the generator's loss takes into account both the magnitude feature loss $\mathscr{L}_{\text{Mag}}$, and the time-domain and frequency-domain features (complex loss) $\mathscr{L}_{\text{R}}$ and $\mathscr{L}_{\text{I}}$. These losses are combined using a linear summation to form the final loss, $\mathscr{L}_{\text{TF}}$, for the model.

$$\begin{aligned} \mathscr{L}_{\text{TF}} &= \alpha \mathscr{L}_{\text{Mag.}} + (1-\alpha)(\mathscr{L}_{\text{R}} + \mathscr{L}_{\text{I}}) \\ \mathscr{L}_{\text{Mag.}} &= \mathbb{E}_{X_m, \hat{X}_m}\left[\|X_m - \hat{X}_m\|^2\right] \\ \mathscr{L}_{\text{R}} + \mathscr{L}_{\text{I}} &= \mathbb{E}_{X_r, \hat{X}_r}\left[\|X_r - \hat{X}_r\|^2\right] + \mathbb{E}_{X_i, \hat{X}_i}\left[\|X_i - \hat{X}_i\|^2\right] \end{aligned} \tag{9}$$

where α is the weight, here we use $\alpha = 0.87$ to train our model.

In the GAN network, the primary goal of the generator is to produce results that are more realistic through adversarial training, in order to deceive the discriminator and achieve better generation performance. The main objective of the discriminator is to distinguish between the generated signals and the real signals. The generator's loss function $\mathscr{L}_{\text{GAN}}$ and the discriminator's loss function $\mathscr{L}_{\text{D}}$ can be defined as:

$$\mathscr{L}_{\text{GAN}} = \mathbb{E}_{\hat{X}_m}\left[\|D(\hat{X}_m) - 1\|^2\right] \tag{10}$$

$$\mathscr{L}_{\text{D}} = \mathbb{E}_{X_m}\left[\|D(X_m) - 1\|^2\right] + \mathbb{E}_{\hat{X}_m}\left[\|D(\hat{X}_m)\|^2\right] \tag{11}$$

where D refers to the discriminator network. Also, we considered the loss in the time domain, which compares the generated respiratory signal with the real respiratory signals:

$$\mathscr{L}_{\text{Time}} = \mathbb{E}_{\text{x},\,\hat{\text{x}}}[\|\text{x} - \hat{\text{x}}\|_1] \tag{12}$$

generally, the loss function of generator can be defined as:

$$\mathscr{L}_{\text{G}} = \gamma_1 \mathscr{L}_{\text{TF}} + \gamma_2 \mathscr{L}_{\text{GAN}} + \gamma_3 \mathscr{L}_{\text{Time}} \tag{13}$$

where, $\gamma_1, \gamma_2, \gamma_3$ are the weights for each respective loss. In this work, we set $\gamma_1 = 0.75$, $\gamma_2 = 0.15$, $\gamma_3 = 0.1$.

3.5. Respiratory quality Index (RQI)

Before calculating the respiratory features, it is essential to evaluate the quality of the generated respiratory signals. In this study, two signal quality indices are used: the Fourier Transform-based ROI_1 and the Autocorrelation-based ROI_2 (Birrenkott et al., 2017; Gazi et al., Jul 2022). The specific evaluation process can be found in (Gazi et al., Jul 2022). To simplify computational complexity, the respiratory signals were downsampled to 4 Hz.

RQI based on FFT (ROI_1): ROI_1 used the power spectrum ($S_i = 1 \ldots 64$) of each 4 Hz, 30 s time-series, as estimated via zero-padded 128-point FFT. The frequency resolution is 1/32. ROI_1 can be defined as follows:

$$ROI_1 = \frac{\hat{S}}{\sum_j s_j} \tag{14}$$

where, $\hat{S}$ can be defined as:

$$\hat{S} = max\left\{ \sum_{k=z-1}^{z} S_k | (z-1), (z) \in [\,222\,] \right\} \tag{15}$$

and where j ∈ [2 22]. The FFT components with indices [2 22] corresponded to the range of frequencies [0.0625 0.6875] Hz. ROI_1 defined the ratio of the maximum power in any 2-component sub-range to the total power in the range of respiratory frequencies.

RQI based on autocorrelation (ROI_2): ROI_2 used the autocorrelation of the respiratory waveform for every sample lag within the physiologically relevant respiratory range (k = 4…64) where each lag denotes a phase shift of $\frac{4}{k}$ Hz.

$$ROI_2 = \max_{k=4\cdots64}\left\{ \frac{\frac{1}{N-1}\sum_{n=1}^{N-k}(s(n) - \bar{x})^*(s(n+k) - \bar{s})}{c_0} \right\} \tag{16}$$

where c_0 is the signal variance, N is the length of the signal, $\bar{s}$ is the signal mean, and k is the signal lag.

3.6. Verification metrics of the model

To validate the model's accuracy, multiple metrics were used to rigorously assess the correlation between the generated respiratory signals and the real respiratory signals.

Root Mean Square Error (RMSE) of the BR: In order to avoid the disturbance caused by the redundant frequency component in calculating BR, the cEDR signal is bandpass filtered (the 0.15–0.4 Hz) according to the normal BR range (9–24 bpm). For the EDR waveforms, the peaks are detected using the method provided by (Khodadad et al., 2018). The BR is acquired from valid peaks for each 30 s segment. The same operation is applied on the corresponding 30 s reference segment. Then, the RMSE of BR for all segments is defined as follows:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(BR_{\text{cEDR}}(k) - BR_{\text{ref}}(k)\right)^2} \tag{17}$$

Where $BR_{\text{cEDR}}(k)$ and $BR_{\text{ref}}(k)$ represent the BR for the k_{th} segment, calculated based on the cEDR and reference signals, respectively. During computation, we apply a 15 s sliding window with median smoothing on both the cEDR and reference signals to enhance signal stability.

Bland–Altman Analysis: Bland–Altman analysis is used to assess the agreement between the generated signals and reference signals by calculating the mean difference and standard deviation of the differences between the two datasets. This method is particularly well-suited for evaluating the consistency between two measurement methods, especially when they share the same units but may exhibit bias or noise.

Pearson correlation coefficient of the BR: Pearson correlation coefficient is used to evaluate the correlation between the generated respiratory signals and the reference signals, ranging from −1 to 1, where values closer to 1 or −1 indicate stronger correlation.

Intraclass Correlation Coefficient (ICC): To evaluate the agreement between the estimated parameters and reference values, this study adopts the ICC(A,1) model, which represents absolute agreement based on a two-way mixed-effects approach. The significance level is set to $\alpha = 0.05$, and the ICC under the null hypothesis is assumed to be $r_0 = 0$, which is used for significance testing and confidence interval estimation. The ICC value is calculated using the following formula:

$$\mathrm{ICC} = \frac{MS_B - MS_W}{MS_B + (k-1)MS_W} \tag{18}$$

where MS_B and MS_W represent the between-subject and within-subject mean squares, respectively, and k is the number of measurements per subject. As a widely used reliability metric, ICC is suitable for assessing the consistency between the respiratory rate measured by the proposed device and the standard reference system.

Retention ratio of respiratory segments: The retention ratio of respiratory segments measures the percentage of valid respiratory signal segments retained during processing. A segment is considered valid if it meets predefined quality criteria, such as signal-to-noise ratio (SNR) thresholds or RQI. The retention ratio provides insight into the reliability and robustness of the model in preserving usable data from the generated signals. High retention ratios indicate that the majority of respiratory segments maintain sufficient quality for further analysis, demonstrating the model's practical applicability.

In respiratory monitoring based on cECG, the RQI is used to screen high-quality signal segments. Factors such as fabric interference, changes in electrode–skin contact impedance, and body position movement can cause fluctuations in cECG signal quality. Particularly during periods of severe signal degradation, the extracted respiratory signal may completely lose physiological significance. Therefore, RQI serves not only as an inclusion criterion for quantitative assessments like Bland-Altman analysis to ensure only reliable data is analysed, but also to calculate the proportion of valid respiratory segments, objectively reflecting the system's usable monitoring duration in real-world applications. This dual application enhances the credibility of the analysis while comprehensively evaluating the system's robustness and practicality in real sleep scenarios.

3.7. Model training and validation strategy

To enhance the generalization capability of the model, we adopted a cross-validation strategy using data from 16 subjects. All subjects were divided into four groups, and in each round, one group (4 subjects) was selected as the test set, while the remaining 12 subjects were further divided into a training set (8 subjects) and a validation set (4 subjects). This process was repeated over four iterations, ensuring that each subject's data served as the test set exactly once. This approach enables comprehensive and fair performance evaluation. During model training, we implement a decaying learning rate schedule where both the generator and discriminator undergo 50 % rate reduction every 10 epochs, with their initial learning rates set to 0.0001 and 0.0002 respectively. A batch size of 16 is used, allowing a balanced quantity of training data for each parameter update, optimizing both training stability and computational efficiency.

In the STFT transformation, we selected a window length of 256 samples to capture at least one full cardiac cycle while ensuring sufficient frequency resolution for distinguishing key waveform components. A step size of 60 samples was chosen to enhance temporal resolution and smoothness.

4. Results

4.1. Typical cECG and generated respiratory signals

Fig. 4 displays typical cECG signals, real respiratory signals, and generated respiratory signals. In Fig. 4(a), the quality of the cECG signal is good, with clear QRS complexes visible. Additionally, it is evident that the generated respiratory signal exhibits a strong waveform correlation with the reference respiratory signal. The peaks and troughs of both signals occur almost simultaneously. In this case, we can also precisely determine the subjects' respiratory states, including the exhalation and inhalation phases. Fig. 4 (d) presents a cECG waveform and its corresponding generated respiratory signal under conditions of poor cardiac coupling. In this case, the cECG waveform lacks distinct QRS complexes, which may result from electrode displacement or the presence of thick coverings (e.g., a blanket) between the body and the electrode. Despite the absence of obvious QRS wave, the cECG signal exhibits strong respiratory coupling due to the body's respiratory movements. Under these conditions, the proposed model still effectively generates a respiratory signal, with the generated waveform (Fig. 4 (f)) showing a notable correlation with the actual respiratory signal (Fig. 4 (e)). Fig. 4 (j) depicts a cECG signal contaminated by significant motion artifacts. When large body movements (such as turning over) occur, the accuracy of the cECG signal significantly deteriorates. In such cases, the cECG signal is unable to reliably capture respiratory information. Similarly, the piezoelectric sensor also experiences substantial noise interference. In this context, the respiratory waveform generated by the TF-GAN model tends to include motion artifact noise, resulting in unreliable respiratory waveforms. This observation highlights that when the cECG signal is compromised by factors such as motion, the generated respiratory signals also lack reliability, further confirming the robustness and limitations of the proposed model.

For the cECG signal shown in Fig. 4 (g) and the generated respiratory waveform in Fig. 4 (i), the cECG demonstrates good performance, with clear QRS complexes and observable respiratory baseline drift. However, the generated respiratory waveform exhibits significant temporal negative correlation with the real respiratory signal shown in Fig. 4(h). Specifically, when the true respiratory signal reaches a peak, the corresponding generated signal shows a trough. This mismatch leads to incorrect respiratory state detection, such as the identification of exhalation and inhalation phases (represented by the blue sections in the Fig. 4).

To explain the observed phenomenon, Equation (3) can be rewritten as:

$$v_{cECG} = k_a v_a - k_b v_b \tag{18}$$

Where, $k_a = \frac{1}{1+\frac{d+\Delta d_a(t)}{Z_{in} S_{\varepsilon r} \varepsilon_0 A_s}}$ and $k_b = \frac{1}{1+\frac{d+\Delta d_b(t)}{Z_{in} S_{\varepsilon r} \varepsilon_0 A_s}}$.

k_a and k_b can be considered as the voltage attenuation coefficients, determined by the electrode–skin interface impedance and the AFE of the cECG signal. Due to variations in sleeping posture, the positions of the electrodes on the subjects differ. To simplify the analysis, we consider two extreme cases of breathing.

When a person lies on both electrodes A and B, the relative displacements of the skin electrodes, $\Delta d_a(t)$ and $\Delta d_b(t)$, vary during respiration. If $\Delta d_a(t) \ll \Delta d_b(t)$, the respiratory modulation is primarily concentrated in $k_a v_a$, resulting in positive modulation. Conversely, when $\Delta d_b(t) \ll \Delta d_a(t)$, the modulation is predominantly focused on $-k_b v_b$, indicating negative modulation.

In actual measurements, the relative changes of $\Delta d_a(t)$ and $\Delta d_b(t)$ are highly complex and uncertain. As a result, the respiratory waveform can

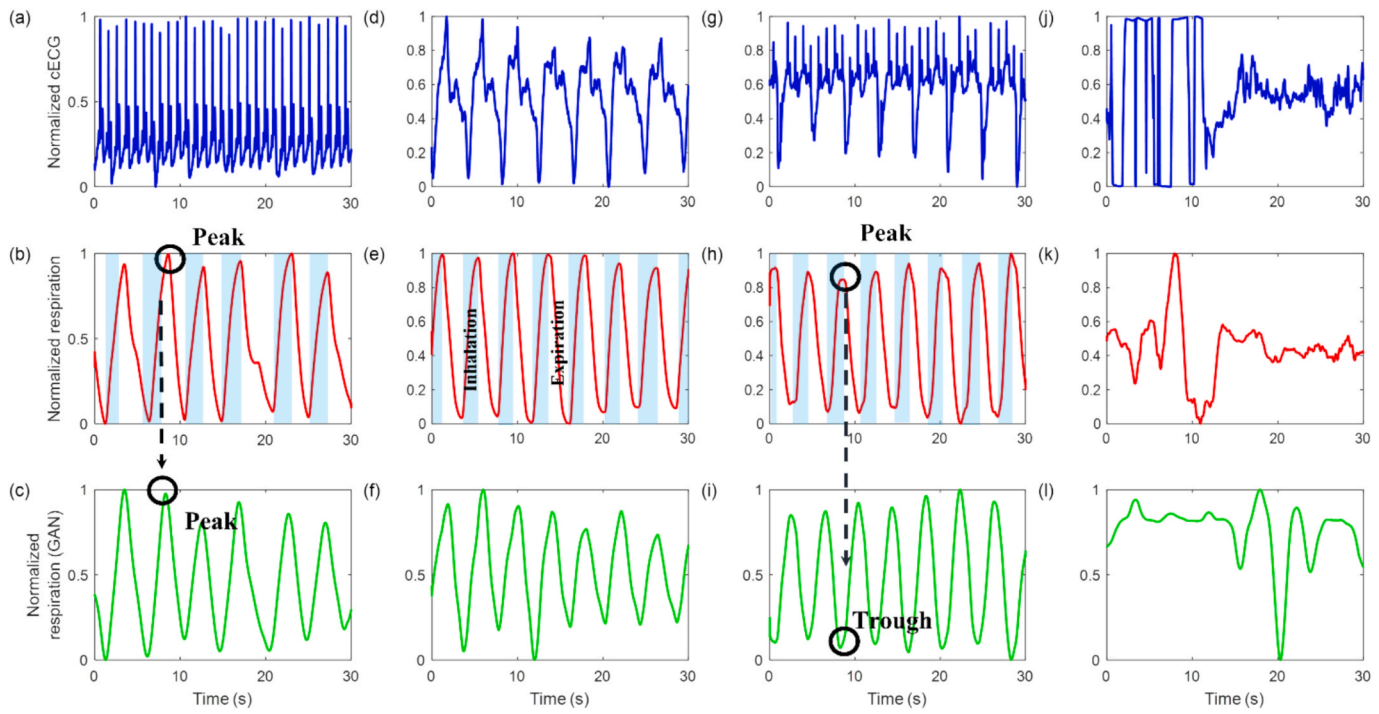


Fig. 4. The typical cECG signal, real respiratory signals, and the respiratory signals generated by the TF-GAN network. (a), (d), (g), (j) display typical cECG signals captured during the actual collection process. (b), (e), (h), (k) show the real-time respiratory signals simultaneously collected using a piezoelectric film sensor. In (b), (e), and (h), the light blue regions represent the inhalation stage, while the rest of the signal corresponds to the expiration stage. (c), (f), (i), and (l) display the respiratory signals generated using the TF-GAN network. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

exhibit both positive and negative correlations. However, this does not affect the calculation of the BR. Nonetheless, for the subsequent identification of inhalation and exhalation phases, additional attention is required.

From the mechanistic perspective, respiratory signal coupling in cECG originates from two primary sources: (1) thoracic volume changes induced by respiration modulate cECG signals through skin-electrode contact impedance (including the relative displacement effects at the interface), and (2) respiratory-induced cardiac axis shifts affect cECG waveforms by altering the relative positions between electrodes and the heart. Respiratory information can be extracted by analyzing cECG waveform morphology, which essentially reflects dynamic variations in cardiac axis relative to electrode positioning. However, in practical sleep monitoring scenarios, postural changes introduce uncertainty in electrode-heart spatial relationships, leading to significant variations in this cardiac axis shift-based respiratory detection mechanism and consequently increasing the complexity of positive/negative respiratory coupling. Experimental results demonstrate that while this mechanistic uncertainty primarily affects respiratory phase identification accuracy, its impact on respiratory rate calculation remains minimal.

This study proposes a deep learning-based method for extracting respiratory signals from cECG measurements. The network effectively captures both QRS complex features and baseline drift features, ensuring model robustness. However, since the respiration extraction process employs deep learning, which has a certain level of inexplicability, it is difficult for the model to distinguish whether thoracic volume or skin-electrode interface impedance contributes more to the extracted respiratory signal during the process. Moreover, considering the characteristics of cECG data, it inherently contains uncertainty, making it difficult for the model to establish a corresponding mapping relationship.

4.2. Evaluation of the generated Respriatory signal

4.2.1. Respiratory waveform correlation assessment

As we observed in the above section, both positive and negative correlations may occur during the generation of respiratory signals. Therefore, we use the pearson correlation coefficient by calculating the correlation coefficients between the generated cEDR waveform and the reference respiratory signal to analyze the distribution of correlations. The Pearson Correlation Coefficient is a statistical measure used to assess the strength of the linear relationship between two variables. The coefficient R ranges from -1 to 1, where $R = 1$ indicates a perfect positive correlation, $R = -1$ indicates a perfect negative correlation, and $R = 0$ suggests no linear relationship between the variables.

Fig. 5 is the distribution of correlation between the generated respiratory signals and the reference signals. The horizontal axis represents the correlation coefficient (ranging from -1 to 1), indicating the degree of similarity between the generated and reference signals, while the vertical axis shows the proportion of each correlation coefficient range. Fig. 6..

It can be observed that 66.4 % of the generated signals have a correlation coefficient greater than 0.5 with the reference signals, indicating a strong positive correlation for approximately 66.4 % of cases. This suggests that, in most instances, the model can generate respiratory waveforms that are highly consistent with the actual signals. Furthermore, a correlation coefficient less than -0.5 is observed in 19.9 % of cases, indicating a negative correlation in approximately 19.9 % of the generated signals. However, since negatively correlated segments can still be valid for calculating BR, their presence does not significantly impact the outcome of BR. The peak in the positive correlation region is at (0.925, 0.200), while the peak in the negative correlation region is around (-0.925, 0.059), highlighting two dominant trends within the data. Further analysis reveals that among all positive correlations, 90.0 % are greater than 0.5; whereas among all negative correlations, 76.7 % are less than -0.5. The difference between these proportions can be

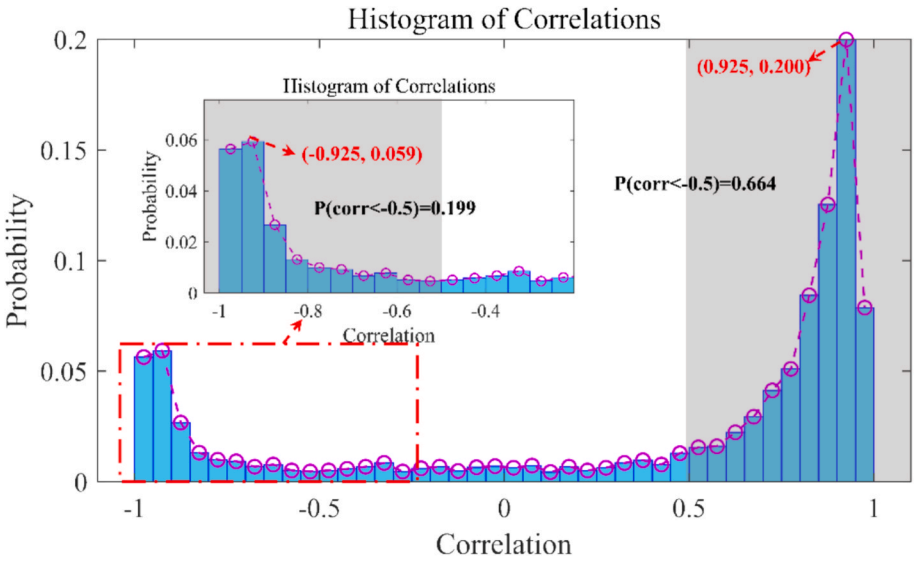


Fig. 5. The correlation distribution between the respiratory waveforms generated by the TF-GAN network and the reference signals.

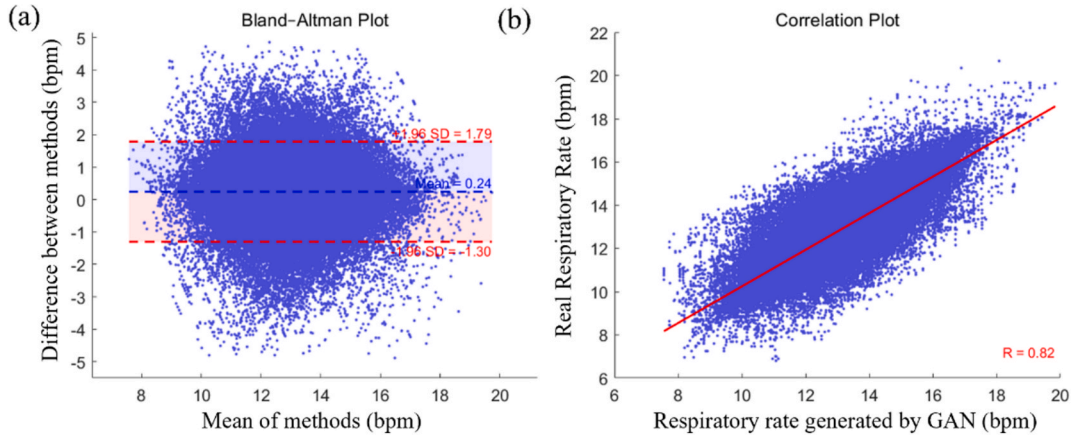


Fig. 6. The accuracy analysis of the respiratory rate calculation based on GAN-generated respiratory waveforms: (a) is the Bland-Altman plot comparing with the true respiratory rate, and (b) is the correlation analysis.

attributed to factors such as individual variability and sleeping posture, indicating that the distribution of correlation values is not symmetric.

Overall, approximately 86.3 % of the signal segments have an absolute correlation coefficient greater than 0.5, further confirming a high level of correlation between the generated and reference signals. In summary, despite the presence of a small proportion of negatively correlated signals, most of the model-generated respiratory signals are highly consistent with the reference signals. Moreover, the negatively correlated signals remain useful for calculating key physiological metrics, such as BR, thus demonstrating the accuracy and applicability of the generated signals produced by the model.

The polarity of respiratory modulation (positive/negative correlation) in cECG signals shows significant variations among individuals. While the overall data demonstrates a predominant trend of positive correlation, the distribution of positive/negative modulation ratios exhibits distinct individual-specific characteristics for each subject. These individual differences primarily stem from subject variability, sleeping postures, and unpredictable changes in skin-electrode impedance, resulting in unpredictable fluctuations in the positive/negative modulation ratio for each participant.

4.2.2. *Bland–Altman analysis for combined patient BR evaluation*

This section presents the results of Bland–Altman analysis for the combined respiratory rates of all patients, which validates the accuracy and.

consistency of our proposed TF-GAN model in generating respiratory signals. Specifically, the mean difference between the generated signals and the reference respiratory signals is 0.241, indicating a minimal systematic bias and demonstrating the model's ability to effectively capture respiratory trends. The standard deviation of the difference is 0.788, showing a high level of consistency in the model's respiratory signal predictions, with relatively low error. The 95 % limits of agreement range from [-1.30, 1.79], indicating that most of the generated signals' errors fall within this range. This is further supported by the finding that 94.72 % of the points lie within the limits of agreement, demonstrating the reliability of the model-generated respiratory signals. Additionally, across all test subjects, the Spearman correlation coefficient between the generated BR and the rate measured by the reference device is 0.82, with a RMSE of 0.97 bpm, indicating a strong positive correlation.

4.3. *Comparison with other EDR methods*

This section presents the Bland–Altman analysis conducted separately for each patient, emphasizing the differences in the mean values and standard deviations across individuals. Table 3 compares the

Table 3
Comparison with Other EDR Methods.

Reference	RMSE of BR (bpm)	Percentage of Points within Limits of Agreement	Pearson Correlation Coefficient of the BR	ICC	Retention Ratio of Respiratory Segments
(Gent et al., 2019)	**1.09 ± 0.09**	95.47 % ± 0.74 %	**0.68 ± 0.10**	**0.59 ± 0.11**	**44.56 % ± 14.88 %**
(Sarkar et al., 2016)	1.27 ± 0.13	95.49 % ± 0.42 %	0.64 ± 0.08	0.51 ± 0.13	29.69 % ± 10.79 %
(Charlton et al., Apr 2016)	1.25 ± 0.14	**95.65 % ± 0.91 %**	0.63 ± 0.12	0.52 ± 0.14	32.11 % ± 12.08 %
(R. Soni, and M. Muniyandi, "Breath Rate Variability: A Novel Measure to Study the Meditation Effects," International Journal of Yoga, vol. 12, 2019)	1.21 ± 0.12	95.10 % ± 0.57 %	0.64 ± 0.13	0.53 ± 0.11	37.18 % ± 12.97 %
Our model	**0.96 ± 0.12**	**94.83 % ± 0.30 %**	**0.74 ± 0.11**	**0.66 ± 0.10**	**67.56 % ± 8.89 %**

performance of respiratory waveform extraction models across different studies utilizing cECG signals. Compared to other models, our proposed model demonstrates superior performance across multiple evaluation metrics. Our model achieves a RMSE of BR of 0.96 ± 0.12 bpm, markedly lower than that of other studies, indicating higher accuracy in BR estimation. Additionally, the proportion of data points within the limits of agreement reaches 94.83 % ± 0.30 %, indicating high consistency and reliability. Furthermore, the BR correlation coefficient is 0.74 ± 0.11, reflecting a strong agreement between the generated signal and the reference signal. Moreover, the ICC reaches 0.66 ± 0.10, further demonstrating good consistency between the estimated and reference respiratory rates.

The model also excels in retaining valid respiratory segments, with an effective segment ratio of 67.56 % ± 8.89 %. Compared to conventional single-feature extraction methods that tend to fail when signal features degrade, our model maintains robust performance under challenging conditions like weakened R-waves or baseline fluctuations by integrating R-wave characteristics, baseline drift, and time–frequency analysis. Since we employed signal quality assessment, other metrics showed relatively moderate improvement as they were calculated only on high-quality segments. The key advantage of our method lies in substantially increasing the number of usable signal segments while maintaining accuracy levels comparable to traditional approaches.

4.4. Robustness to noise

To evaluate the robustness of the signal generation algorithm under different noise conditions. This was achieved by superimposing Gaussian noise onto the original noncontact cECG signals. Since the original cECG signals inherently contain a certain level of noise, additional Gaussian noise was introduced to create more challenging noise conditions for analysis. Specifically, the Relative-noise SNR is defined as the ratio of the baseline signal power to the artificially added noise power:

$$\text{Relative-noise SNR} = 10\log_{10}\left(\frac{P_{cECG}}{P_{add}}\right) \tag{19}$$

Where, P_{cECG} refers to the power of the cECG signal, P_{add} refers to the power of the added noise.

The target SNRs were set to 0 dB, 7 dB, and 15 dB to simulate varying levels of noise interference. The noise power was adjusted to achieve the desired SNR levels, ensuring the signal was tested under controlled conditions. For example, at an SNR of 0 dB, the added noise resulted in the signal power being approximately equal to the noise power (a ratio of 1:1). The test signals were based on originally acquired cECG data, with the added noise being zero-mean Gaussian white noise with specific variances. Under these different noise conditions, we assessed the performance of the respiratory generation algorithm and evaluated its robustness through comparative analysis.

The selection of these SNR levels, with the choice of 15 dB attenuation, reflects common recording conditions in practical non-contact cECG measurements. Previous studies have explored the impact of fabric thickness and moisture regain on signal quality. Results show that under standard atmospheric conditions, the SNR of cECG signals acquired using cotton fabrics with a thickness of less than 1 mm typically ranges from −5 dB to 15 dB (Xiao et al., 2022). In sleep scenarios, localized humidity increases due to body heat and perspiration, leading to higher fabric moisture regain. Under these conditions, the SNR generally falls within the range of 0 dB to 15 dB. Therefore, a 15 dB relative SNR attenuation represents a testing condition that closely mimics the actual environment in sleep-related cECG monitoring, better reflecting the signal challenges encountered in real-world applications.

Table 4 presents a comparative analysis of our proposed algorithm against other methods under varying noise conditions. Our algorithm demonstrates significant superiority across different noise levels, particularly in key metrics such as RMSE of BR, Pearson correlation coefficient, and respiratory segment retention ratio.

Under noise conditions of 0 dB, 7 dB, and 15 dB, the RMSE values of our algorithm are 1.05 ± 0.21 bpm, 0.97 ± 0.15 bpm, and 0.97 ± 0.12 bpm, respectively. These values are notably lower than those of other methods (1.16 ± 0.14 bpm, 1.12 ± 0.13 bpm, and 1.09 ± 0.10 bpm, respectively), showcasing higher accuracy in estimating breathing rates. Furthermore, our algorithm excels in the Pearson correlation coefficient between the estimated and reference breathing rates, achieving 0.67 ± 0.10, 0.71 ± 0.11, and 0.73 ± 0.09 at 0 dB, 7 dB, and 15 dB, respectively, indicating robust performance even under high-noise conditions.

In terms of the retention ratio of respiratory segments, our algorithm outperforms others significantly, achieving 61.47 % ± 10.89 % at 0 dB and increasing to 66.97 % ± 9.72 % at 15 dB. In contrast, the retention ratios of reference methods remain consistently lower. Additionally, as noise levels increase (from 15 dB to 0 dB), the retention ratio of our method decreases by only 5.50 %, compared to a reduction of 6.35 % to 11.49 % observed in other methods.

This advantage is attributable to our algorithm's ability to rely on features beyond the QRS waveform of the cECG signal, enabling accurate respiratory waveform generation even when QRS features are less prominent.

5. Discussion

This paper presents a TF-GAN-based model for generating respiratory signals from cECG signals. The experimental results demonstrate that the model can accurately capture the characteristics of actual respiratory waveforms in most cases. Even under conditions of poor cECG signal quality, where QRS complexes are difficult to identify, the model can still generate signals with prominent respiratory features, highlighting its applicability and robustness in various signal coupling environments. However, when the cECG signal is contaminated by significant motion artifacts, the generated respiratory signal is also affected, containing substantial motion-related noise. This is consistent with real-world conditions; however, for more refined respiratory disorder assessments, effective noise removal is essential.

Currently, the RQI used in this study are relatively simple and do not

Table 4
Comparison of cEDR Methods Across Different Noise Levels in cECG Signals.

Reference	Noise Levels	RMSE of BR (bpm)	Percentage of Points within Limits of Agreement	Pearson Correlation Coefficient of the BR	ICC	Retention Ratio of Respiratory Segments
(Gent et al., 2019)	0 dB	**1.16 ± 0.14**	95.21 % ± 0.60 %	**0.67 ± 0.13**	**0.55 ± 0.16**	**32.95 % ± 9.78 %**
	7 dB	**1.12 ± 0.13**	95.19 % ± 0.30 %	**0.68 ± 0.11**	**0.60 ± 0.13**	**42.29 % ± 12.72 %**
	15 dB	**1.09 ± 0.10**	95.53 % ± 0.30 %	**0.68 ± 0.09**	**0.59 ± 0.11**	**43.34 % ± 14.25 %**
(Sarkar et al., 2016)	0 dB	1.30 ± 0.14	94.97 % ± 0.61 %	0.60 ± 0.16	0.47 ± 0.14	21.87 % ± 6.57 %
	7 dB	1.29 ± 0.18	**95.65 % ± 0.96 %**	0.61 ± 0.10	0.50 ± 0.13	28.22 % ± 8.72 %
	15 dB	1.27 ± 0.13	**95.48 % ± 0.60 %**	0.64 ± 0.08	0.51 ± 0.16	28.44 % ± 10.14 %
(Charlton et al., Apr 2016)	0 dB	1.33 ± 0.18	95.39 % ± 0.78 %	0.58 ± 0.14	0.44 ± 0.12	23.91 % ± 6.35 %
	7 dB	1.31 ± 0.22	95.25 % ± 0.46 %	0.61 ± 0.09	0.46 ± 0.15	30.01 % ± 8.31 %
	15 dB	1.25 ± 0.15	95.28 % ± 0.39 %	0.62 ± 0.12	0.50 ± 0.14	30.26 % ± 10.62 %
(R. Soni, and M. Muniyandi, "Breath Rate Variability: A Novel Measure to Study the Meditation Effects," International Journal of Yoga, vol. 12, 2019)	0 dB	1.24 ± 0.07	**95.41 % ± 0.51 %**	0.63 ± 0.09	0.52 ± 0.08	25.09 % ± 8.27 %
	7 dB	1.23 ± 0.10	94.96 % ± 0.58 %	0.64 ± 0.13	0.54 ± 0.15	35.37 % ± 11.97 %
	15 dB	1.21 ± 0.12	95.11 % ± 0.34 %	0.64 ± 0.14	0.53 ± 0.12	36.58 % ± 12.38 %
Our model	0 dB	**1.05 ± 0.21**	**94.58 % ± 0.65 %**	**0.67 ± 0.10**	**0.60 ± 0.12**	**61.47 % ± 10.89 %**
	7 dB	**0.97 ± 0.15**	**94.98 % ± 0.54 %**	**0.71 ± 0.11**	**0.64 ± 0.09**	**62.79 % ± 8.98 %**
	15 dB	**0.97 ± 0.12**	**94.74 % ± 0.63 %**	**0.73 ± 0.09**	**0.67 ± 0.10**	**66.97 % ± 9.72 %**

fully evaluate the fine quality of the generated signals. Future work should focus on developing more advanced and complex respiratory quality assessment models to meet the high accuracy and detail requirements of respiratory detection. This would help improve the model's application in fine-grained respiratory monitoring tasks, particularly in clinical respiratory status monitoring and daily health management, ensuring higher reliability and accuracy. Additionally, signal quality assessment should comprehensively consider the quality of cECG. By combining the quality assessments of both cECG signals and the generated respiratory signals, the overall trustworthiness of the signals can be more accurately judged, thereby enhancing the accuracy and practicality of the generated signals. This multi-level signal quality evaluation system would help ensure the reliability and precision of the generated signals in practical applications.

Moreover, the current model still faces challenges in respiratory phase identification. Although the generated signals exhibit high accuracy in BR calculation, some generated signals may exhibit phase inversion when compared to reference signals, i.e., issues of positive and negative correlation. This means that the peaks and troughs of the generated signals may not perfectly match the real signals, affecting the accurate classification of respiratory phases. As such, the current model requires further improvement in respiratory phase recognition, particularly in ensuring the accurate distinction between inhalation and exhalation. This issue should also be carefully considered in practical applications to prevent negative impacts on the accuracy of respiratory status monitoring.

To further enhance the performance of the model, future research should focus on the following areas. First, expanding the training dataset to increase sample size will improve the model's generalization ability and robustness. More diverse samples will help the model better accommodate physiological differences among individuals and various practical scenarios, thus enhancing its broad applicability. These improvements will provide stronger support for the model in more detailed respiratory monitoring tasks, enabling more reliable use in clinical respiratory status monitoring and daily health management.

6. Conclusion

Respiration monitoring and analysis play a crucial role in detecting sleep-related disorders. This paper explores a novel approach to breath monitoring using cEDR to record respiratory waveforms during sleep. We propose a method based on the Time and Frequency domain-GAN Network (TF-GAN) to generate respiratory signals from cECG. the TF-GAN model shows excellent performance in generating accurate respiratory signals from contactless ECG (cECG) data. The experimental results show that the generated respiratory signals exhibit a high correlation with the reference signals. Specifically, 86.3 % of the signals have a absolute waveform correlation coefficient greater than 0.5, indicating good reproduction of real breathing waveforms. Our proposed model demonstrates superior performance in respiratory signal extraction, achieving a low RMSE of 0.96 ± 0.12 bpm and a high agreement rate of 94.83 % ± 0.30 % within the Bland–Altman limits. Additionally, the model maintains an effective respiratory segment ratio of 67.56 % ± 8.89 %, even under poor cECG signal conditions, showcasing its robustness and reliability.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was funded by the National Natural Science Foundation of China (62171123,62211530112), the National Key Research and Development Program of China (2023YFC3603600), the Natural Science Foundation of Jiangsu Province (BK20200371, BK20192004), and

in part by the CSC scholarship program (202306090058). Maarten De Vos and Christos Chatzichristos are supported by the Flemish Government (AI Research Program) and FWO Research Project: 'Artificial Intelligence (AI) for data-driven personalised medicine' (G0C9623N).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.eswa.2025.128360.

Data availability

The authors do not have permission to share data.

References

K. Dong, L. Zhao, C. Zou, Z. Cai, C. Yan, Y. Li, J. Li, and C. Liu, "A Novel ECG-Derived Respiration Method Combining Frequency-Domain Feature and Interacting Multiple Model Smoother," *IEEE Trans Biomed Eng,* vol. 70, no. 3, pp. 888-898, Sep 7, 2022.

Jin, H., Tao, X., Dong, S. R., Qin, Y. H., Yu, L. Y., Luo, J. K., & Deen, M. J. (Nov 2017). Flexible surface acoustic wave respiration sensor for monitoring obstructive sleep apnea syndrome. *Journal of Micromechanics and Microengineering, 27*(11), 7.

W. Do, R. Russell, C. Wheeler, M. Lockwood, M. De Vos, I. Pavord, and M. Bafadhel, "Performance of Contactless Respiratory Rate Monitoring by Albus Home(TM), an Automated System for Nocturnal Monitoring at Home: A Validation Study," *Sensors (Basel),* vol. 22, no. 19, Sep 21, 2022.

Chiu, Y.-Y., Lin, W.-Y., Wang, H.-Y., Huang, S.-B., & Wu, M.-H. (2013). Development of a piezoelectric polyvinylidene fluoride (PVDF) polymer-based sensor patch for simultaneous heartbeat and respiration monitoring. *Sensors and Actuators A: Physical, 189*, 328–334.

L. Jin, Z. K. Liu, M. Altintas, Y. Zheng, Z. C. Liu, S. R. Yao, Y. Y. Fan, and Y. Li, "Wearable Piezoelectric Airflow Transducers for Human Respiratory and Metabolic Monitoring," *Acs Sensors*, pp. 12, 2022 Jul, 2022.

Song, C. X., Wang, J. G., Zhao, P. C., Zhang, Y. P., & Zhang, Q. (Aug 2024). A novel method for chronic obstructive pulmonary disease detection using comprehensive respiratory impedance and electrocardiogram-derived respiration. *Measurement, 236*, 11.

Garbino, A., Blue, R. S., Pattarini, J. M., Law, J., & Clark, J. B. (Feb 2014). Physiological Monitoring and Analysis of a Manned Stratospheric Balloon Test Program. *Aviation Space and Environmental Medicine, 85*(2), 177–182.

Dong, K., Zhao, L., Cai, Z., Li, Y., Li, J., & Liu, C. (Aug 2021). An integrated framework for evaluation on typical ECG-derived respiration waveform extraction and respiration. *Comput Biol Med, 135*, Article 104593.

Thomas, R. J., Mietus, J. E., Peng, C. K., & Goldberger, A. L. (Sep 2005). An electrocardiogram-based technique to assess cardiopulmonary coupling during sleep. *Sleep, 28*(9), 1151–1161.

Uguz, D. U., Canbaz, Z. T., Antink, C. H., Luken, M., & Leonhardt, S. (2022). A Novel Sensor Design for Amplitude Modulated Measurement of Capacitive ECG. *IEEE Transactions on Instrumentation and Measurement, 71*, 1–10.

Wang, K., Yao, Y., Lin, R., Cheng, A., Xu, Y., & Xu, L. (Mar 2023). A Capacitive Electrocardiography System With Dedicated Noise-Cancellation Algorithms for Morphological Analysis. *IEEE Trans Biomed Eng, 70*(3), 1086–1094.

Leicht, L., Eilebrecht, B., Weyer, S., Leonhardt, S., & Teichmann, D. (Apr 2017). Closed-Loop Control of Humidification for Artifact Reduction in Capacitive ECG Measurements. *IEEE Trans Biomed Circuits Syst, 11*(2), 300–313.

Wang, T.-W., Zhang, H., & Lin, S.-F. (2020). Influence of Capacitive Coupling on High-Fidelity Non-Contact ECG Measurement. *IEEE Sensors Journal, 20*(16), 9265–9273.

Orphanidou, C., Fleming, S., Shah, S. A., & Tarassenko, L. (Jan 2013). Data fusion for estimating respiratory rate from a single-lead ECG. *Biomedical Signal Processing and Control, 8*(1), 98–105.

Bailón, R., Sörnmo, L., & Laguna, P. (2006). ECG-derived respiratory frequency estimation. *Advanced methods and tools for ECG data analysis, 1*, 215–244.

E. Helfenbein, R. Firoozabadi, S. Chien, E. Carlson, and S. Babaeizadeh, "Development of three methods for extracting respiration from the surface ECG: A review," *Journal of Electrocardiology,* vol. 47, no. 6, pp. 819-825, Nov-Dec, 2014.

Xie, J. L., Fonseca, P., Van Dijk, J., Long, X., & Overeem, S. (Jul 2023). The Use of Respiratory Effort Improves an ECG-Based Deep Learning Algorithm to Assess Sleep-Disordered Breathing. *Diagnostics, 13*(13), 15.

Tyagi, P. K., & Agrawal, D. (Feb 2023). Automatic detection of sleep apnea from single-lead ECG signal using enhanced-deep belief network model. *Biomedical Signal Processing and Control, 80*, 8.

Hu, S., Liu, J., Yang, C., Wang, A., Li, K., & Liu, W. (2023). Semi-supervised learning for low-cost personalized obstructive sleep apnea detection using unsupervised deep learning and single-lead electrocardiogram. *IEEE Journal of Biomedical and Health Informatics*.

Lindberg, L.-G., Ugnell, H., & Öberg, P. (1992). Monitoring of respiratory and heart rates using a fibre-optic sensor. *Medical and Biological Engineering and Computing, 30*, 533–537.

W. Yi, and K. Park, "Derivation of respiration from ECG measured without subject's awareness using wavelet transform." pp. 130-131.

Sadr, N., & de Chazal, P. (Feb 2019). A comparison of three ECG-derived respiration methods for sleep apnoea detection. *Biomedical Physics & Engineering Express, 5*(2), 15.

Lázaro, J., Alcaine, A., Romero, D., Gil, E., Laguna, P., Pueyo, E., & Bailón, R. (2014). Electrocardiogram derived respiratory rate from QRS slopes and R-wave angle. *Annals of biomedical engineering, 42*, 2072–2083.

Uguz, D. U., Tufan, T. B., Uzun, A., Leonhardt, S., & Hoog Antink, C. (2020). Physiological Motion Artifacts in Capacitive ECG: Ballistocardiographic Impedance Distortions. *IEEE Transactions on Instrumentation and Measurement, 69*(6), 3297–3307.

Fleming, S. G., & Tarassenko, L. (2007). A comparison of signal processing techniques for the extraction of breathing rate from the photoplethysmogram. *Int. J. Biol. Med. Sci, 2*(4), 232–236.

Shelley, K. H., Awad, A. A., Stout, R. G., & Silverman, D. G. (2006). The use of joint time frequency analysis to quantify the effect of ventilation on the pulse oximeter waveform. *Journal of clinical monitoring and computing, 20*, 81–87.

Garde, A., Karlen, W., Ansermino, J. M., & Dumont, G. A. (2014). Estimating respiratory and heart rates from the correntropy spectral density of the photoplethysmogram. *PloS one, 9*(1), 1–11.

Xiao, Z. J., Ma, C. Y., Xing, Y. T., Yang, C. X., Hao, M. L., Li, J. Q., & Liu, C. Y. (2023). Atrial Fibrillation Monitoring Based on Noncontact Capacitive ECG Using an Integrated Microhumidity Fabric Electrode-Sheet Sensing Scheme. *Ieee Transactions on Instrumentation and Measurement, 72*, 11.

Quintana, D. S., Elstad, M., Kaufmann, T., Brandt, C. L., Haatveit, B., Haram, M., Nerhus, M., Westlye, L. T., & Andreassen, O. A. (2016). Resting-state high-frequency heart rate variability is related to respiratory frequency in individuals with severe mental illness but not healthy controls. *Scientific Reports, 6*(1), 37212.

Abdulatif, S., Cao, R., & Yang, B. (2024). CMGAN: Conformer-Based Metric-GAN for Monaural Speech Enhancement. *IEEE/ACM Transactions on Audio, Speech, and Language Processing, 32*, 2477–2493.

A. Pandey, and D. Wang, "Densely connected neural network with dilated convolutions for real-time speech enhancement in the time domain." pp. 6629-6633.

Chan, T. K., & Chin, C. S. (2021). Multi-Branch Convolutional Macaron net for Sound Event Detection. *IEEE/ACM Transactions on Audio, Speech, and Language Processing, 29*, 2972–2985.

A. Gulati, J. Qin, C. C. Chiu, N. Parmar, Y. Zhang, J. Yu, W. Han, S. Wang, Z. Zhang, and Y. Wu, "Conformer: Convolution-augmented Transformer for Speech Recognition," *arXiv preprint arXiv:2005.08100*, 2020.

P. Ramachandran, B. Zoph, and Q. V. Le, "Searching for activation functions," *arXiv preprint arXiv:1710.05941*, 2017.

Li, A., Zheng, C., Zhang, L., & Li, X. (2021). "Glance and Gaze. *A Collaborative Learning Framework for Single-channel Speech Enhancement,"*.

Birrenkott, D. A., Pimentel, M. A., Watkinson, P. J., & Clifton, D. A. (2017). A robust fusion model for estimating respiratory rate from photoplethysmography and electrocardiography. *IEEE Transactions on Biomedical Engineering, 65*(9), 2033–2041.

Gazi, A. H., Jung, H., Kimball, J. P., & Inan, O. T. (Jul 2022). Improving Respiratory Timing Estimation Using Quality Indexing and Electrocardiogram-Derived Respiration. *Annu Int Conf IEEE Eng Med Biol Soc, 2022*, 3249–3252.

Khodadad, D., Nordebo, S., Muller, B., Waldman, A., Yerworth, R., Becher, T., Frerichs, I., Sophocleos, L., Kaam, A. V., & Miedema, M. (2018). "Optimized breath detection algorithm in Electrical Impedance Tomography," *IEEE Transactions on Biomedical Engineering,* vol. *PP, 99*, 460–467.

Gent, P. V., Farah, H., Nes, N. V., & Arem, B. V. (2019). HeartPy: A novel heart rate algorithm for the analysis of noisy signals. *Transportation Research Part F: Traffic Psychology and Behaviour, 66*, 368–378.

Sarkar, S., Bhattacherjee, S., & Pal, S. (2016). Extraction of respiration signal from ECG for respiratory rate estimation. *Michael Faraday Iet International Summit*.

Charlton, P. H., Bonnici, T., Tarassenko, L., Clifton, D. A., Beale, R., & Watkinson, P. J. (Apr 2016). An assessment of algorithms to estimate respiratory rate from the electrocardiogram and photoplethysmogram. *Physiological Measurement, 37*(4), 610–626.

R. Soni, and M. Muniyandi, "Breath Rate Variability: A Novel Measure to Study the Meditation Effects," *International Journal of Yoga,* vol. 12, 2019.

Xiao, Z., Xing, Y., Li, Y., Yang, C., Li, J., & Liu, C. (2022). Effect of Cotton Fabric Moisture Regain and Thickness on Signal Quality of Noncontact Capacitive Coupling ECG. *IEEE Transactions on Instrumentation and Measurement, 71*, 1–12.