sensors

MDPI

Article

A Variational Autoencoder Network for Heartbeat Signal Reconstruction and Breathing-State Classification Using 24 GHz Radar Sensors

Homa Arab [1,*], Nima Suzandeh [2], Serioja O. Tatu [2] and Steven Dufour [1]

[1] Department of Mathematics and Industrial Engineering, Polytechnique Montreal, Montreal, QC H3T 1J4, Canada; steven.dufour@polymtl.ca
[2] Institut National de la Recherche Scientifique (INRS), Energie Materiaux et Telecommunications, Montreal, QC H5A 1K6, Canada; nima.souzandeh@inrs.ca (N.S.); serioja.ovidiu.tatu@inrs.ca (S.O.T.)
* Correspondence: homa.arab1983@gmail.com

Abstract

Wireless radar sensors can remotely detect vital signs. This study investigates the wireless monitoring of physiological signals induced by respiration and heartbeat to classify breathing states. Two 24 GHz radar sensors are used. Noise, harmonics, and intermodulation products affect the measured physiological signals and can make heartbeat identification and analysis challenging. Deep-learning techniques have recently been used to denoise and classify signals in wireless communications. Variational autoencoders (VAEs) have also been used for signal and image processing. Multilayer VAEs support sequential learning by optimizing generative-model parameters and reducing reconstruction error. This study investigates the separation of measured heartbeat and respiratory signals and the reduction in noise and the effects of limited random body movements using VAEs under controlled conditions. Two radars are positioned in front of and behind a human subject to record vital signals simultaneously. The separated and denoised heartbeat signals are input to a ResNet network to classify normal and fast breathing states. The proposed method achieves 99.7% accuracy on the reported experimental dataset.

Keywords: artificial intelligence; convolutional neural network (CNN); deep learning; denoising; Doppler frequency; machine learning; microwave; modulator; receiver; recurrent neural network (RNN); ResNet; signal classification; signal processing; variational autoencoder (VAE)

Check for updates

Academic Editor: Dipankar Mitra

Received: 4 September 2026
Revised: 21 October 2026
Accepted: 23 October 2026
Published: 9 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Radars can accurately detect a target's Doppler frequency and wirelessly estimate its range [1,2]. Microwave radar sensors have proven to be accurate contactless sensors that can detect small movements and vibrations in the micrometer range [3]. They are used in applications such as non-contact vital-sign detection, sleep monitoring, human-fall detection, smart surveillance, and security [4,5]. They also offer greater flexibility than wearable sensors such as electrocardiogram (ECG) monitors, conventional photoplethysmography (PPG) sensors, and piezoresistive sensors.

Multiple-radar systems are attracting growing attention because they can simultaneously record different signals [6,7]. Recording several waveforms from a target enables the extraction of additional information for target detection and localization. In this work, we

use two 24 GHz radar sensors to measure heartbeat and respiratory signals and classify normal and fast breathing states.

The proposed radar systems transmit orthogonal signals from two transmitter (TX) antennas, and the reflected signals are captured by the antennas of each receiver (RX). This setup makes it possible to collect long-term data from two receivers and detect small movements of one or more objects. Extracting the desired information from the received signal is an important and challenging task because of the various types of noise and interference found in RX signals, the nonstationary nature of the target, the existence of multiple targets, and the various RX antenna signals.

Various methods have been proposed to suppress noise and interference. Low-pass filters are commonly used to suppress noise in RF signals. Their implementation can be complex, and their parameters must be tuned to avoid attenuating low-frequency information. Machine-learning-based signal denoising has recently become popular in wireless communication systems, where these methods have shown strong performance in denoising and classification [8].

Recent studies applying deep-learning methods to heartbeat detection and classification are listed in Table 1. Most use ECG signals as input and convolutional neural networks (CNNs) or long short-term memory (LSTM) networks to classify raw time-domain signals or their scalograms. This study focuses on reconstructing heartbeat signals from physiological signals using VAEs. CNNs are also used for the final breathing-state classification. The VAE is a generative model used for signal and image processing, data augmentation, and generative modeling [9,10]. VAEs have the ability to provide meaningful latent representations of the input data produced by the encoder, and to generate new output signals from the compressed latent space via the decoder [11].

Table 1. A survey of machine learning neural network applications for heartbeat signal reconstruction and classification.

Ref.	Instrument	Machine Learning Technique	Accuracy	Application
[12]	Electrocardiogram (ECG)	MPCNN	96.4%	Inter-Patient ECG Classification
[13]	Electrocardiogram (ECG)	CNN	98.45%	Heartbeat ECG Classification
[14]	Electrocardiogram (ECG)	ELM-LRF and BLSTM	99.32% and 97.15%	Arrhythmia Detection
[15]	Electrocardiogram (ECG)	ResNet-101	99.75%	Heartbeat ECG Classification
[16]	MIT-BIH Database	CNN/LSTM	98%	Heartbeat Detection
[17]	Doppler radar 24 GHz	LSTM	NA	Heartbeat Signal Reconstruction
[18]	CW Doppler Radar 2.4 GHz	CNN/LSTM	86.6%	Sleep Stage Estimation
[19]	Doppler radar 5.8 GHz	ConceFT algorithm	NA	Heart-Rate Detection
[20]	CW Doppler Radar 2.4 GHz	SVM, LDA, and NB	88%, 86%, and 80%	Vital Signs Monitoring
Proposed	Two CW Doppler Radars, 24 GHz	Variational Autoencoders (VAEs); ResNet Convolutional Networks	99.7%	Heartbeat Signal Reconstruction and Breathing-State Classification

The contributions of this article can be summarized as follows:

- Two 24 GHz radar sensors are used to measure the Doppler frequencies of a human subject's physiological signals simultaneously. To the best of the authors' knowledge, this is the first study to employ two radar sensors for this specific application. The sensors provide measurements from two directions at the same time. The radar measurement accuracy is assessed using metronome-based measurements. The mean frequency error is ± 0.01 Hz, and the mean range error is ±10 micron [1].
- A VAE-based method is used to extract a heartbeat signal from a physiological signal and remove noise and interference from the heartbeat waveform. VAEs are powerful deep-learning techniques that have been used for signal and image processing. To the best of the authors' knowledge, this is the first application of this method to heartbeat reconstruction and breathing-state classification.
- The reconstructed and denoised signals are input to a ResNet-50 neural network to classify normal and fast breathing states. The network achieves an accuracy of 99.7%.
- A shifting-based augmentation technique is used to double the dataset by shifting each sample by 1 ms along the time axis. The experimental results show that combining this augmentation technique with the ResNet-50 and VAE networks achieves 99.7% accuracy. The proposed technique may facilitate the extraction of essential physiological measures, such as breathing rate, pulse rate, body temperature, and blood pressure, from the denoised heartbeat signal.

2. Mathematical and Simulation Analyses

As illustrated in Figure 1, two identical commercial 24 GHz CW Doppler radars are positioned in front of and behind a target. The block diagram and operating parameters of the radar are shown in Figure 1 and Table 2. The radar operates at a center frequency of 24 GHz, has a bandwidth of 1 GHz, and an equivalent isotropically radiated power (EIRP) of 30 dBm. It includes a low-noise microwave amplifier (LNA) to amplify the input signal. A balun at the receiver input converts the antenna's unbalanced signal into the balanced differential signal required by the receiver or mixer. It also provides impedance matching and improves common-mode noise rejection. This compact, embedded radar can be integrated into various devices. A 24 GHz microwave radar transmits a CW signal that can be written as

$$T(t) = A_{\text{TX}} \cos\big(2\pi f t + \theta(t)\big), \tag{1}$$

where $\theta(t)$ represents the phase shift created by the active components, A_{TX} is the amplitude level of the transmitted signal, and f is the center frequency. The target's chest moves due to the heartbeat and the expansion and contraction of the lungs. Therefore, the signal received by the front radar contains information about both heartbeat and breathing rates. The radar behind the target, however, receives a reflected signal that contains only heartbeat information.

The reflected signal from the chest of the human target, denoted by $R(t)$, can be written as

$$R(t) = A_{\text{RX}} \cos\big(2\pi f (t - t_d) + \theta(t - t_d)\big), \tag{2}$$

where A_{RX} is the received-signal amplitude, which depends on the transmitter and receiver antenna gains, the radar cross section, the target distance, and the carrier wavelength. The propagation delay, t_d, is given by

$$t_d = \frac{2d(t)}{c}. \tag{3}$$

The time-varying distance between the radar and the target is

$$d(t) = d_0 + x(t), \tag{4}$$

where $x(t)$ represents periodic chest movement, and d_0 is the distance between the radar and the target. Substituting Equations (3) and (4) into Equation (2), the received signal can be rewritten as

$$R(t) = A_{\text{RX}} \cos\Big(2\pi f t - \big(\tfrac{4\pi d_0 + 4\pi x(t)}{\lambda}\big) + \theta\big(t - \tfrac{2d(t)}{c}\big)\Big). \tag{5}$$

The baseband in-phase (I) and quadrature-phase (Q) signals are given by:

$$\begin{aligned} I(t) &= A_I \cos\Big(\big(\tfrac{4\pi d_0 + 4\pi x(t)}{\lambda}\big) + \theta\big(t - \tfrac{2d(t)}{c}\big)\Big); \\ Q(t) &= A_Q \sin\Big(\big(\tfrac{4\pi d_0 + 4\pi x(t)}{\lambda}\big) + \theta\big(t - \tfrac{2d(t)}{c}\big)\Big), \end{aligned} \tag{6}$$

where A_I and A_Q are the amplitudes of the I and Q signals, respectively. The residual phase noise $\theta(t) - \theta\big(t - 2d(t)/c\big)$ is negligible because the sensor is coherent. The TX and local oscillator (LO) signals share a source and phase noise, which can therefore cancel. The chest displacement $x(t)$ can be modeled as a function of the breathing and heartbeat rates:

$$\begin{aligned} x(t) &= x_{\text{r}}(t) + x_{\text{h}}(t); \\ x(t) &= m_{\text{r}} \cos(2\pi f_{\text{r}} t) + m_{\text{h}} \cos(2\pi f_{\text{h}} t), \end{aligned} \tag{7}$$

where $x_{\text{r}}(t)$ and $x_{\text{h}}(t)$ are the chest displacements due to respiration and heartbeat, respectively. The quantities m_{r} and m_{h} are their displacement magnitudes, while f_{r} and f_{h} are their frequencies. Substituting Equation (7) into Equation (6), the baseband in-phase (I) and quadrature-phase (Q) signals can be written as

$$\begin{aligned} I(t) &= A_I \cos\Big[\tfrac{4\pi x_{\text{r}}(t)}{\lambda} + \tfrac{4\pi x_{\text{h}}(t)}{\lambda} + \tfrac{4\pi d_0}{\lambda}\Big]; \\ Q(t) &= A_Q \sin\Big[\tfrac{4\pi x_{\text{r}}(t)}{\lambda} + \tfrac{4\pi x_{\text{h}}(t)}{\lambda} + \tfrac{4\pi d_0}{\lambda}\Big]. \end{aligned} \tag{8}$$

To analyze the harmonics and frequency components of the received radar signal, we use a Fourier-series expansion involving Bessel functions $J_i(x)$ of order i. Equation (9) shows that respiratory harmonics and intermodulation products between the respiratory and cardiac signals can affect detection accuracy. The heartbeat peak amplitude depends on both m_{r} and m_{h}; it corresponds to the terms $(k = 0, l = \pm 1)$.

$$\begin{aligned} I(t) = \sum_{k=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} & J_k\Big(\tfrac{4\pi m_{\text{r}}}{\lambda}\Big) J_l\Big(\tfrac{4\pi m_{\text{h}}}{\lambda}\Big) \\ & \cos\big(2\pi(k f_{\text{r}} + l f_{\text{h}})t\big) \cos\Big(\tfrac{4\pi d_0}{\lambda}\Big); \\ Q(t) = \sum_{k=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} & J_k\Big(\tfrac{4\pi m_{\text{r}}}{\lambda}\Big) J_l\Big(\tfrac{4\pi m_{\text{h}}}{\lambda}\Big) \\ & \sin\big(2\pi(k f_{\text{r}} + l f_{\text{h}})t\big) \sin\Big(\tfrac{4\pi d_0}{\lambda}\Big). \end{aligned} \tag{9}$$

As mentioned earlier, the reflected signal from behind the target contains only heartbeat information. Therefore, the baseband in-phase I and quadrature-phase Q can be written as:

$$\begin{aligned} I(t) &= \sum_{l=-\infty}^{+\infty} J_l\Big(\tfrac{4\pi m_{\text{h}}}{\lambda}\Big) \cos(2\pi l f_{\text{h}} t) \cos\Big(\tfrac{4\pi d_0}{\lambda}\Big); \\ Q(t) &= \sum_{l=-\infty}^{+\infty} J_l\Big(\tfrac{4\pi m_{\text{h}}}{\lambda}\Big) \sin(2\pi l f_{\text{h}} t) \sin\Big(\tfrac{4\pi d_0}{\lambda}\Big). \end{aligned} \tag{10}$$

In the experimental setup shown in Figure 1, a Tektronix DPO7054 oscilloscope displays the amplitude, phase, and period of the received baseband I and Q signals. An Agilent DSO-X 2024A oscilloscope provides a second display for the Lissajous curve in the I/Q plane. The direction of motion along a circle or spiral in this plane (clockwise or counterclockwise) depends on the direction of the target's chest movement. A DC source provides a tuning voltage between 0 and 5 V to adjust the frequency across the 1 GHz bandwidth, from 23.5 GHz to 24.5 GHz. A TFC-3000 digital universal frequency counter measures and

displays the transmitted-signal frequency; the displayed value is multiplied by 8192. The setup uses simple, low-cost components.

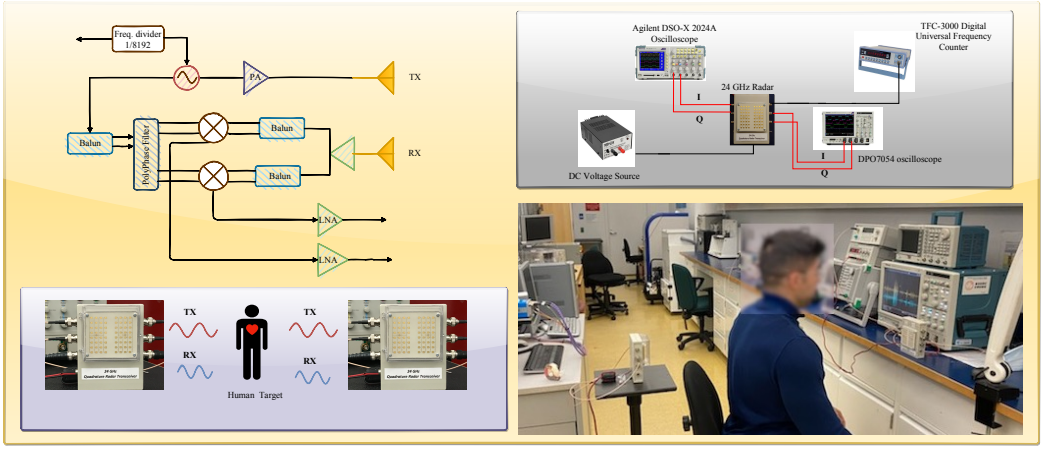


Figure 1. Block diagram and experimental setup for measuring heartbeat and respiration using two 24 GHz CW radars positioned in front of and behind a human subject.

Table 2. Characteristics of the proposed radar [21].

Characteristic	**Value**
Center frequency	24 GHz
Frequency bandwidth deviation	1 GHz (23.5 to 24.5 GHz)
Horizontal & vertical −3 dB beamwidth	24° & 16°
Side lobe level (*H*-plane)	−29 dB
Side lobe level (*E*-plane)	−11 dB
TX & RX antenna gain	16 dBi
First IF LNA	20 dB
Second IF LNA	14 dB
Ambient temperature range	[−25 °C, 85 °C]
Storage temperature range	[−40 °C, 105 °C]

Figure 2 shows the captured in-phase baseband signals in the time and frequency domains for several conditions: normal breathing, fast breathing, and breath-holding for 30 s. The waveforms were recorded by radar sensors positioned in front of and behind a human subject. Their spectra show the respiration frequency and its first three harmonics, intermodulation components, and the heartbeat frequency. In the next section, we describe a deep generative network for extracting relevant features from the recorded signals. A ResNet-50 CNN is used to classify normal and fast breathing states.

The present evaluation was conducted under controlled laboratory conditions and included only limited random body shifts. Minor changes in posture or target position can alter the radar-target distance, angle, phase, and signal amplitude, which may affect VAE reconstruction and heartbeat–respiration separation. The VAE can therefore reduce the effects of the limited movements observed in this setup, but the current study does not provide quantitative results for fully unconstrained movement. A systematic evaluation using predefined positional shifts and reconstruction, signal-to-noise, and classification metrics is left for future work.

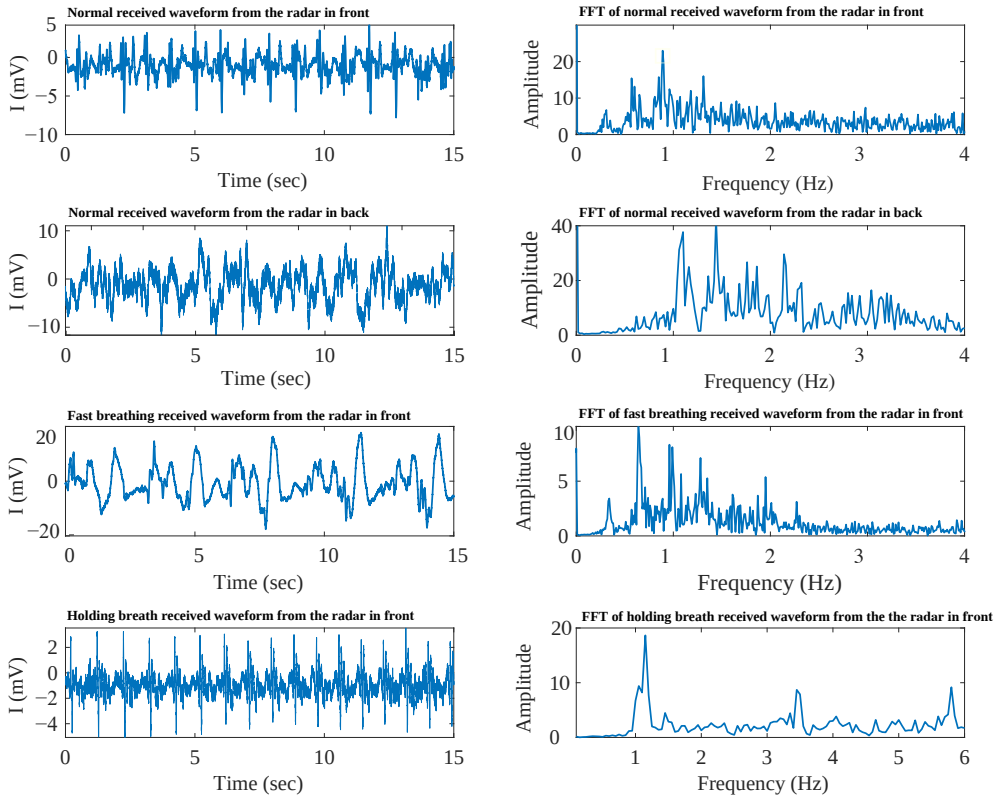


Figure 2. Waveform and spectrum of received signal in various states.

3. Variational Autoencoders (VAEs)

A variational autoencoder (VAE) is a deep generative model introduced by Diederik P. Kingma and Max Welling in 2013 [22]. It has an architecture similar to that of a standard autoencoder, but differs in its mathematical formulation and the distributions used in its latent space. These properties reduce discontinuities in the latent representation and improve the model's ability to generate outputs. VAEs have been used in image generation [23], text generation [24], learning basic visual concepts [25], and speech-spectrogram modeling [26].

A VAE has three components: a probabilistic encoder, a sampling unit, and a decoder. The basic architecture of a VAE is shown in Figure 3.

The encoder, or recognition model, comprises convolutional networks that take the input ($\mathbf{x}$) and map it to a lower-dimensional latent representation (z). The marginal likelihood is

$$p_\theta(x) = \int p_\theta(z) p_\theta(x|z)\, dz.$$

The encoder approximates the true posterior distribution, $p_\theta(z|x)$, using $q_\phi(z|x)$. This approximate posterior is assumed to be Gaussian, and ϕ denotes the encoder's variational parameters. The encoder outputs the mean and standard deviation of $q_\phi(z|x)$. The decoder, denoted by $p_\theta(x|z)$, attempts to reconstruct the original input from the latent representation z. A sampling unit draws the latent vector z from the approximate posterior as follows:

$$z \sim q_\phi(z|x) = N(\mu, \sigma) \tag{11}$$

where μ is the mean vector and σ is the standard-deviation vector.

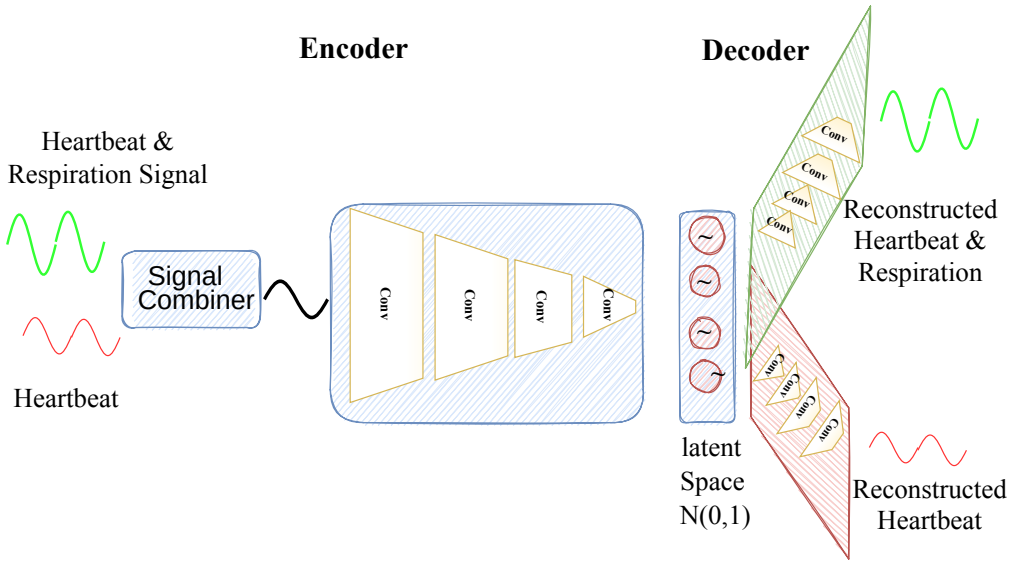


Figure 3. Architecture of the proposed VAE network.

The variational parameters ϕ and generative-distribution parameters θ are learned by training the VAE. The loss function to be minimized has two components: a reconstruction loss and a regularization loss, given by

The VAE architecture used in this work is shown in Figure 4. The VAE receives two time-domain input signals. Before augmentation, the raw VAE dataset contained 15,000 samples. It was divided into 10,000 training samples (66.7%), 2500 validation samples (16.7%), and 2500 testing samples (16.7%). All samples came from a single participant ($n = 1$), and no k-fold cross-validation was performed. Thus, these results demonstrate VAE feasibility under the considered experimental conditions, but do not establish subject-independent generalizability. The first input signal contains heartbeat and respiratory components recorded by the front radar. The second contains the heartbeat waveform recorded by the radar behind the target. Each input column contains 100,000 points, which are reduced to 10,000 by averaging each consecutive set of 10 points. This factor-of-10 reduction was selected to balance signal preservation and computational cost. Averaging consecutive samples attenuates rapid sample-to-sample fluctuations before decimation, thereby reducing high-frequency noise and the memory and computational requirements of the VAE. Because the heartbeat and respiratory components are lower-frequency physiological signals relative to the original acquisition rate, the procedure is intended to retain their temporal waveforms, including the small amplitude variations associated with heartbeat micro-movements, while suppressing unwanted high-frequency fluctuations. The 10,000 points are divided into windows of 256 points with a stride of 8. The batch size is 60. As shown in Figure 4, the encoder uses four convolutional layers with 32, 64, 128, and 256 channels, each with a kernel size of 3. After encoding, the signal is flattened into a one-dimensional feature vector used to compute the mean and variance of the latent distribution. These parameters are used to sample the latent space; the resulting samples are passed to the decoder to reconstruct the heartbeat and respiratory signals.

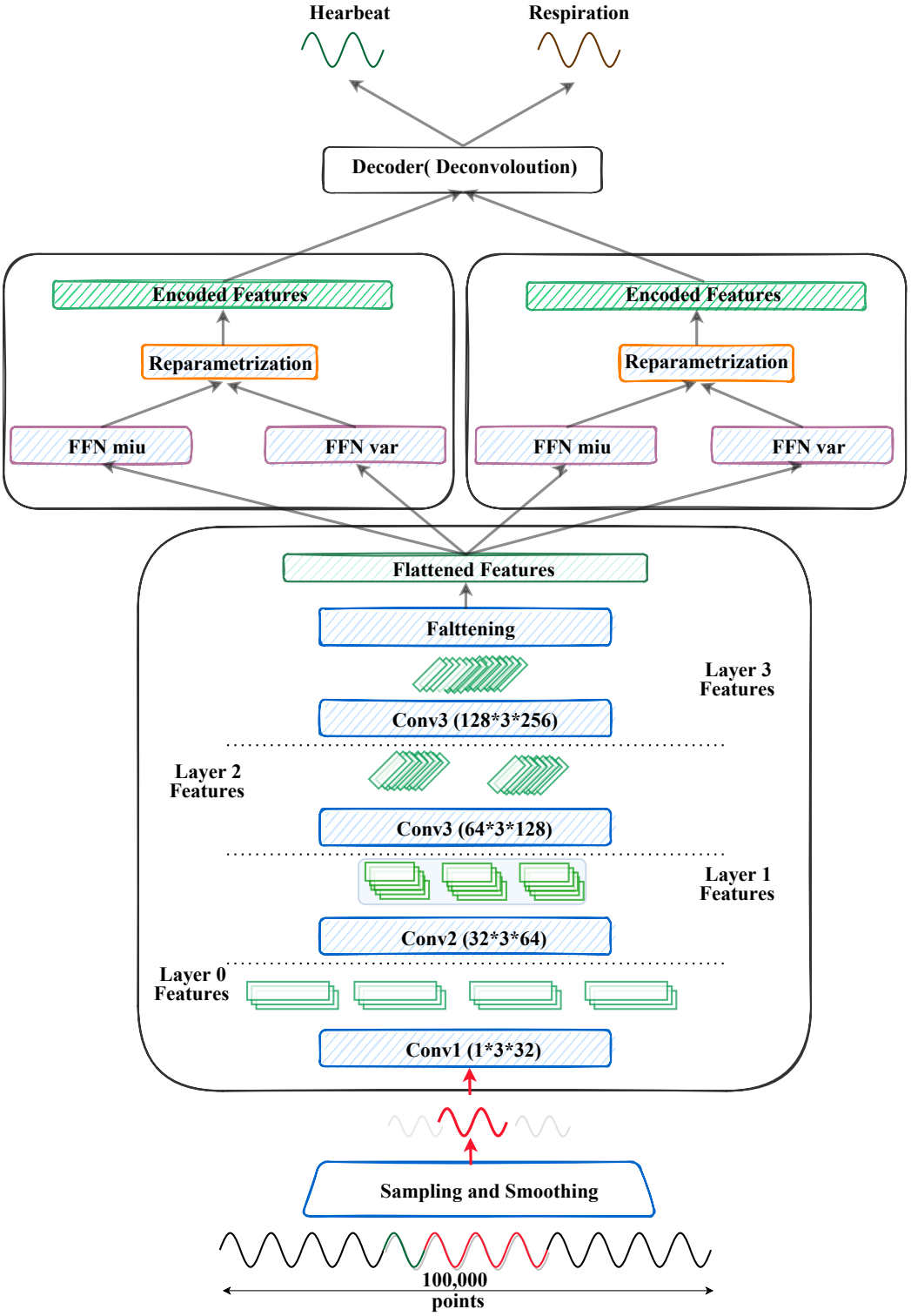


Figure 4. Block diagram Block diagram and key details of the VAE used in this study.

$$\mathcal{L}(\theta, \phi) = -\mathbb{E}_{z \sim q_\phi(z|x)}(\log p(\tilde{x}|z)) + D_{KL}(q_\phi(z|x) \| p_\theta(z)) \tag{12}$$

where $\mathcal{L}$ denotes the VAE loss associated with the encoder and decoder parameters. The first term is the reconstruction term, which measures the discrepancy between the input and its reconstruction. The second term, $D_{KL}(\cdot)$, is the Kullback–Leibler (KL) divergence

between the approximate posterior and the latent-variable prior. Minimizing this term regularizes the approximate posterior toward the prior distribution.

$$D_{KL}(q_\phi(z|x)\|p_\theta(z)) = \frac{1}{2}\sum_{j=1}^{J}\left(\mu_j^2 + \sigma_j^2 - 1 - \log\sigma_j^2\right). \quad (13)$$

Further details on the derivation of the VAE loss are provided in [9,10]. All convolutional layers use the LeakyReLU activation function, and 30% dropout is used to help prevent overfitting. Each layer also includes a batch-normalization layer to reduce training time and improve stability. The encoder output is flattened into a feature vector used to compute the mean and variance of the latent space. The decoder comprises four transposed-convolution layers and one convolutional layer. The Adam optimizer, with $\beta_1 = 0.9$ and $\beta_2 = 0.999$, is used for training.

A learning-rate scheduler with a step size of 40 epochs and a gamma of 0.9 is used during training. Figure 5a shows the signal recorded by the front radar. Figure 5b,c show the VAE output signals corresponding to respiration and heartbeat, respectively. The denoised and separated heartbeat signal is then input to a ResNet network for classification of normal and fast breathing states.

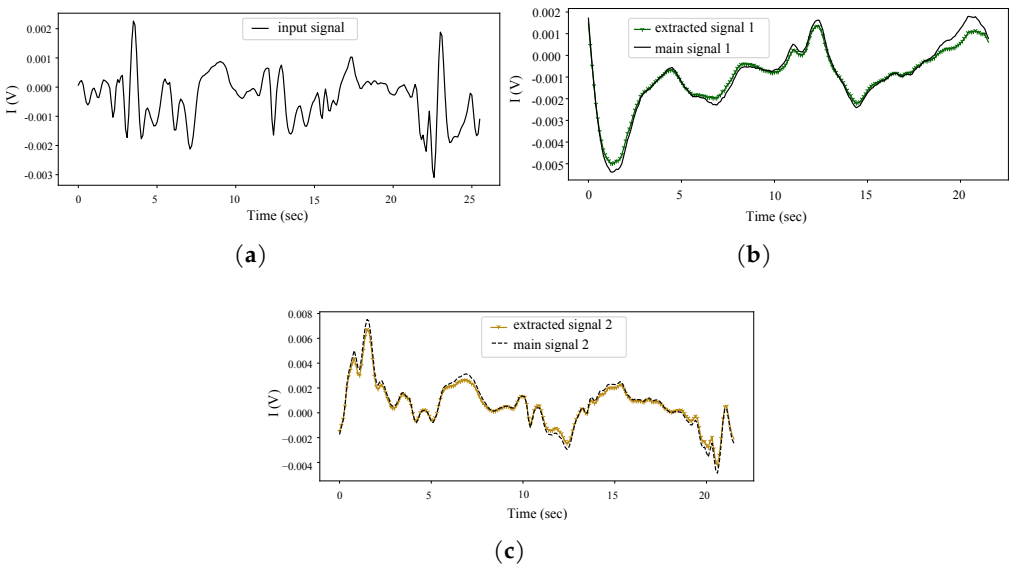


Figure 5. Input and output signals of the VAE network. The input contains heartbeat and respiratory components; the outputs represent respiration (signal 1) and heartbeat (signal 2). (**a**) The VAE input signal includes heartbeat and respiratory signals; (**b**) The VAE output signal represents the respiratory signal; (**c**) The VAE output signal represents the heartbeat signal.

ResNet Convolutional Networks

Convolutional neural networks (CNNs) are deep-learning methods for classification and feature extraction from many types of data. Their use in image and signal processing has grown substantially in recent decades. To improve classification and estimation accuracy, some researchers use CNNs with more layers, known as deep convolutional neural networks (DCNNs). Adding convolutional layers can improve classification accuracy until the network saturates; beyond that point, accuracy may decrease, and overfitting and training error may increase [27]. Vanishing gradients can also arise in DCNNs and may prevent the weights from converging during backpropagation. He et al. introduced residual neural networks (ResNets) to address vanishing-gradient and saturation problems [28]. The ResNet-50 architecture used in this work is shown in Figure 6, and its bottleneck block is shown in Figure 7.

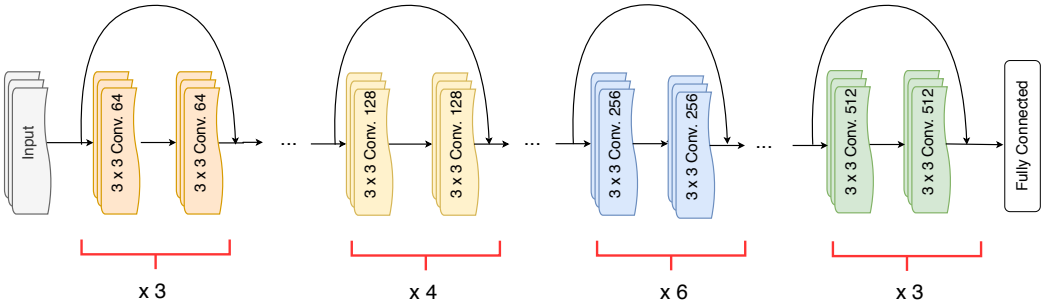


Figure 6. Architecture diagram of ResNet-50.

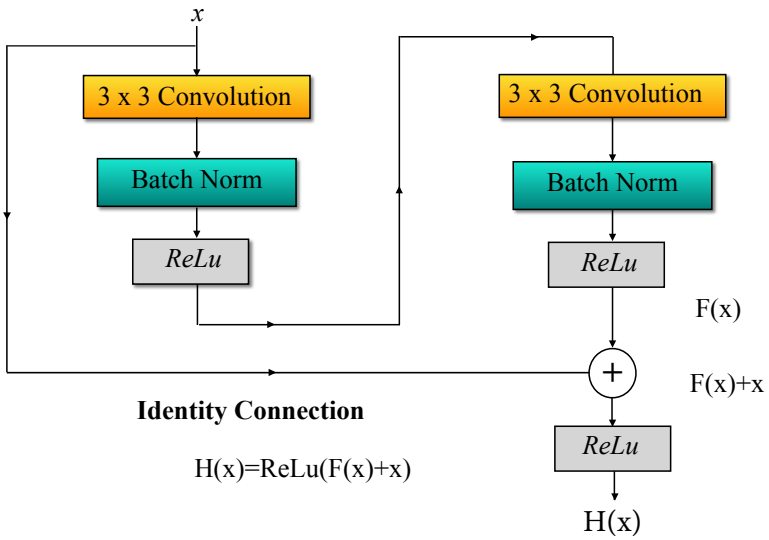


Figure 7. A bottleneck block of a residual network with two layers.

ResNet uses identity connections between layers to pass input data directly to deeper layers.

Skip connections allow information to pass through deeper layers and help the network learn identity mappings. The residual mapping, $F(x)$, represents the difference between the input and output of a residual block. The residual mapping is added to the block input to produce the block output, $H(x)$. In training the residual network, the main focus is finding and minimizing $F(x)$. After finding $F(x)$, the deeper layers can also produce extra feature maps.

Due to laboratory-related limitations, measurements were conducted under two breathing conditions: normal and fast breathing. The input dataset comprises 250 baseband I/Q signals across these two classes. All recorded signals are from a single participant ($n = 1$). Age, gender, health status, positioning variability, and subject-to-subject variability were not assessed; therefore, the results are not generalizable to other individuals or populations. The ResNet-50 is applied to the I/Q signals collected by the 24 GHz radar sensors positioned in front of and behind the subject, after VAE-based separation of the respiratory components. Data augmentation is used to expand the training dataset and help reduce overfitting. In this work, a shifting-based augmentation technique doubles the dataset by shifting each sample by 1 ms along the time axis.

The Adam optimizer with a learning rate of 0.001 and a cross-entropy loss function was used for classification. The ResNet-50 achieved a classification accuracy of 99.7 % after 100 training epochs. A conventional CNN achieved an accuracy of 96.5 % under the reported experimental conditions.

The proposed VAE ResNet-50 pipeline achieved a 3.2 percentage-point improvement in breathing-state classification accuracy over the conventional CNN under the same reported

experimental conditions. The VAE is designed to learn an adaptive nonlinear mapping from the paired front- and rear-radar signals, enabling noise and motion-artifact suppression while reconstructing the heartbeat and respiratory components. In contrast, conventional low-pass, band-pass, or wavelet filtering is primarily frequency-based and can attenuate desired vital-sign components when they overlap spectrally with respiratory harmonics, motion artifacts, or intermodulation products. Learning-based denoisers can also provide adaptive suppression, but their performance depends on the training objective, architecture, and availability of representative paired noisy and clean signals. The current accuracy comparison evaluates the end-to-end classification pipeline and does not isolate the VAE contribution.

Figure 8 presents the confusion matrices for the two networks. They show the numbers of false positives, false negatives, true positives, and true negatives for normal and fast breathing. In the displayed results, the ResNet-50 correctly classifies both classes with no errors. Future work will use the extracted heartbeat signal to estimate additional physiological measures.

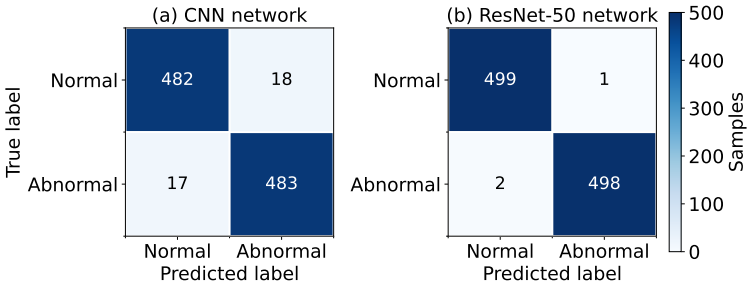


Figure 8. Confusion matrices for the system using (**a**) a CNN and (**b**) ResNet-50.

4. Conclusions

This study investigated the wireless monitoring of physiological signals induced by respiration and heartbeat using two microwave radar sensors. A variational autoencoder (VAE) network was proposed to separate the heartbeat signal from the respiratory signal and denoise the heartbeat component. Under controlled conditions, the method reduced the effects of noise and limited random body movement in the measured heartbeat signal. A ResNet-50 network was then used to classify the denoised heartbeat signal. The proposed method achieved an overall accuracy of 99.7% in classifying normal versus fast breathing. Further work may investigate estimating physiological measures, including heart rate, breathing rate, temperature, and blood pressure, from the denoised heartbeat signal.

Author Contributions: Conceptualization, H.A., S.O.T. and S.D.; methodology, H.A. and S.O.T.; software, H.A. and S.D.; validation, H.A. and N.S.; formal analysis, H.A.; investigation, H.A. and N.S.; resources, S.O.T.; data curation, H.A.; writing—original draft preparation, H.A.; writing—review and editing, N.S., S.O.T. and S.D.; supervision, S.O.T. and S.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable. This study in non-invasive and will not cause any harm to the human subjects involved.

Data Availability Statement: Data is available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Arab, H.; Dufour, S.; Moldovan, E.; Akyel, C.; Tatu, S.O. Accurate and Robust CW-LFM Radar Sensor: Transceiver Front-End Design and Implementation. *IEEE Sens. J.* **2019**, *19*, 1943–1950. [CrossRef]
2. Arab, H.; Chioukh, L.; Ardakani, M.D.; Dufour, S.; Tatu, S.O. Early-Stage Detection of Melanoma Skin Cancer Using Contactless Millimeter-Wave Sensors. *IEEE Sens. J.* **2020**, *20*, 7310–7317. [CrossRef]
3. Arab, H.; Dufour, S.; Moldovan, E.; Akyel, C.; Tatu, S.O. A 77 GHz Six-Port Sensor for Accurate Near Field Displacement and Doppler Measurements. *Sensors* **2018**, *18*, 2565. [CrossRef]
4. Pauli, M.; Göttel, B.; Scherr, S.; Bhutani, A.; Ayhan, S.; Winkler, W.; Zwick, T. Miniaturized Millimeter-Wave Radar Sensor for High-Accuracy Applications. *IEEE Trans. Microw. Theory Tech.* **2017**, *65*, 1707–1715. [CrossRef]
5. Arab, H.; Ghaffari, I.; Chioukh, L.; Tatu, S.O.; Dufour, S. A Convolutional Neural Network for Human Motion Recognition and Classification Using a Millimeter-Wave Doppler Radar. *IEEE Sens. J.* **2022**, *22*, 4494–4502 . [CrossRef]
6. Frischen, A.; Hasch, J.; Waldschmidt, C. A Cooperative MIMO Radar Network Using Highly Integrated FMCW Radar Sensors. *IEEE Trans. Microw. Theory Tech.* **2017**, *65*, 1355–1366. [CrossRef]
7. Han, K.; Hong, S. Detection and Localization of Multiple Humans Based on Curve Length of I/Q Signal Trajectory Using MIMO FMCW Radar. *IEEE Microw. Wirel. Compon. Lett.* **2021**, *31*, 413–416. [CrossRef]
8. Arab, H.; Ghaffari, I.; Evina, R.M.; Tatu, S.O.; Dufour, S. A Hybrid LSTM-ResNet Deep Neural Network for Noise Reduction and Classification of V-Band Receiver Signals. *IEEE Access* **2022**, *10*, 14797–14806. [CrossRef]
9. Liu, Z.-S.; Siu, W.-C.; Chan, Y.-L. Photo-Realistic Image Super-Resolution via Variational Autoencoders. *IEEE Trans. Circuits Syst. Video Technol.* **2021**, *31*, 1351–1365. [CrossRef]
10. Cristovao, P.; Nakada, H.; Tanimura, Y.; Asoh, H. Generating In-Between Images Through Learned Latent Space Representation Using Variational Autoencoders. *IEEE Access* **2020**, *8*, 149456–149467. [CrossRef]
11. Yan, S.; Smith, J.S.; Lu, W.; Zhang, B. Abnormal Event Detection From Videos Using a Two-Stream Recurrent Variational Autoencoder. *IEEE Trans. Cogn. Dev. Syst.* **2020**, *12*, 30–42. [CrossRef]
12. Niu, J.; Tang, Y.; Sun, Z.; Zhang, W. Inter-Patient ECG Classification With Symbolic Representations and Multi-Perspective Convolutional Neural Networks. *IEEE J. Biomed. Health Inform.* **2020**, *24*, 1321–1332. [CrossRef]
13. Maghawry, E.; Gharib, T.F.; Ismail, R.; Zaki, M.J. An Efficient Heartbeats Classifier Based on Optimizing Convolutional Neural Network Model. *IEEE Access* **2021**, *9*, 153266–153275. [CrossRef]
14. Qiao, F.; Li, B.; Zhang, Y.; Guo, H.; Li, W.; Zhou, S. A Fast and Accurate Recognition of ECG Signals Based on ELM-LRF and BLSTM Algorithm. *IEEE Access* **2020**, *8*, 71189–71198. [CrossRef]
15. Zhang, Y.; Li, J.; Wei, S.; Zhou, F.; Li, D. Heartbeats Classification Using Hybrid Time-Frequency Analysis and Transfer Learning Based on ResNet. *IEEE J. Biomed. Health Inform.* **2021**, *25*, 4175–4184. [CrossRef]
16. Schwob, M.R.; Dempsey, A.; Zhan, F.; Zhan, J.; Mehmood, A. Robust Multimodal Heartbeat Detection Using Hybrid Neural Networks. *IEEE Access* **2020**, *8*, 82201–82214. [CrossRef]
17. Yamamoto, K.; Ohtsuki, T. Non-Contact Heartbeat Detection by Heartbeat Signal Reconstruction Based on Spectrogram Analysis With Convolutional LSTM. *IEEE Access* **2020**, *8*, 123603–123613. [CrossRef]
18. Hong, H.; Zhang, L.; Gu, C.; Li, Y.; Zhou, G.; Zhu, X. Noncontact Sleep Stage Estimation Using a CW Doppler Radar. *IEEE J. Emerg. Sel. Top. Circuits Syst.* **2018**, *8*, 260–270. [CrossRef]
19. Saluja, J.; Casanova, J.; Lin, J. A Supervised Machine Learning Algorithm for Heart-Rate Detection Using Doppler Motion-Sensing Radar. *IEEE J. Electromagn. Rf Microwaves Med. Biol.* **2020**, *4*, 45–51. [CrossRef]
20. Hanifi, K.; Karsligil, M.E. Elderly Fall Detection with Vital Signs Monitoring Using CW Doppler Radar. *IEEE Sens. J.* **2021**, *21*, 16969–16978. [CrossRef]
21. Milandr Integrated Circuit Design Center, 24 GHz Doppler Radar Sensor ("Overview 24" Transceiver). Available online: https://www.milandr.ru/ (accessed on 21 October 2026).
22. Kingma, D.P.; Welling, M. Auto-encoding variational Bayes. *arXiv* **2014**, arXiv:1312.6114.
23. Yan, X.; Yang, J.; Sohn, K.; Lee, H. Attribute2Image: Conditional image generation from visual attributes. In *Proceedings of the Computer Vision–ECCV 2016, Amsterdam, The Netherlands, 8–16 October 2016*; Springer: Cham, Switzerland, 2016; pp. 776–791. [CrossRef]
24. Wang, W.; Gan, Z.; Xu, H.; Zhang, R.; Wang, G.; Shen, D.; Chen, C.; Carin, L. Topic-guided variational autoencoders for text generation. *arXiv* **2019**, arXiv:1903.07137.
25. Higgins, I.; Matthey, L.; Pal, A.; Burgess, C.; Glorot, X.; Botvinick, M.; Mohamed, S.; Lerchner, A. Beta-VAE: Learning basic visual concepts with a constrained variational framework. In Proceedings of the International Conference on Learning Representations, Toulon, France, 24–26 April 2017.
26. Nugraha, A.A.; Sekiguchi, K.; Yoshii, K. A flow-based deep latent variable model for speech spectrogram modeling and enhancement. *IEEE/ACM Trans. Audio Speech Lang. Process.* **2020**, *28*, 1104–1117. [CrossRef]

27. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. In *Proceedings of the Advances in Neural Information Processing Systems*; MIT Press: Cambridge, MA, USA, 2012; pp. 1097–1105.
28. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; IEEE: New York, NY, USA, 2016; pp. 770–778.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.