Generating Breathing Patterns In Real-Time: Low-Latency Respiratory Phase Tracking From 25Hz PPG

Ian Karman*[0009−0006−1226−7193], Yue Sun*[0009−0008−5528−2566], Rahil Soroushmojdehi[0000−0003−0319−0885], Jose A. Silva[0009−0001−8883−5488], and Mostafa 'Neo' Mohsenvand**[0000−0001−8114−3617]

BrainCo Inc, Somerville, MA, USA
ian.karman@brainco.tech, yue.sun@brainco.tech,
rahil.soroushmojdehi@brainco.tech, jose.silva@brainco.tech,
neo.mohsenvand@brainco.tech

Abstract. This study presents a low-latency, real-time breathing cycle tracking system utilizing a conditional Generative Adversarial Network (GAN) with Wasserstein loss, with a low-powered, low sample rate Photoplethysmography (PPG) sensor. The aim is to provide a clinically accurate respiratory tool capable of tracking and visualizing the breathing cycle and rate in real-time for at-home and general ambulatory applications. To detect breathing activity in real-time, we used a wearable headband with a 25 Hz PPG sensor and an inductive respiratory sensor as ground truth. To meet the real-time and low latency constraints, the inputs were processed in 1-second windows. Signal processing and machine learning techniques were explored, and the proposed GAN-based method with Wasserstein loss and gradient penalty, outperformed others in accurately tracking the ground-truth breathing curve. Leveraging the GAN-generated breathing curve, a peak-detection algorithm calculated the respiratory rate (RR) with an average mean absolute error (MAE) of 1.47 breaths per minute (bpm) across 10 test subjects, comparable to high-sampling rate PPG literature (1 bpm), but with the advantage of 5 times faster real-time monitoring. The GAN-generated respiratory signal from a low-sampling rate wearable PPG sensor demonstrates potential as a viable alternative to traditional respiratory monitoring systems. This system offers valuable breathing monitoring, useful in various applications such as pain management.

Keywords: Machine Learning · Respiratory Tracking · Photoplethysmography · Generative Adversarial Networks · Sequence-to-Sequence · Real-Time Monitoring · Low-Latency · Respiratory Rate · Breathing Patterns.

---

* These authors contributed equally to this work.
** Corresponding Author

1 Introduction

The monitoring of patients, both in hospital settings and at home, has significantly benefited from the advancements in wearable devices. These devices provide the ability to continuously and non-invasively track various physiological parameters, enabling timely interventions and improving patient care. Among the vital signs, monitoring breathing in real-time plays a crucial role in assessing respiratory function and detecting respiratory abnormalities [24]. Real-time monitoring of breathing using wearable devices plays a vital role in pain management applications, promoting mindful breathing exercises, and offering valuable insights into a patient's respiratory patterns, respiratory rate (RR), and overall respiratory health. These devices allow individuals to track their breathing patterns in real-time, facilitating better pain control techniques and enhancing overall well-being. By providing immediate feedback on breathing cycles, wearable devices enable individuals to engage in mindful breathing exercises, helping to reduce stress, anxiety, and promote relaxation [10,24]. The convenience and mobility offered by these devices make them ideal for at-home use, empowering individuals to take an active role in managing their pain and engaging in mindfulness practices outside of traditional healthcare settings. Real-time breathing monitoring has promising applications beyond healthcare settings, extending its benefits to athletes, swimmers, and singers. Athletes can optimize their breathing patterns for enhanced performance and injury prevention. Swimmers benefit from improved stroke efficiency and breath control, while singers can refine their breath support and vocal projection. These applications empower individuals in various domains to achieve their best performance and elevate their skills [24].

Several previous works have explored the provision of feedback for breathing control, each employing different techniques and technologies [10]. For instance, Breeze is a mobile application that utilizes a smartphone's microphone to continuously detect breathing phases, subsequently triggering a gamified biofeedback-guided breathing training [20,38]. BreatheBuddy, on the other hand, employs the low-power accelerometer in earbuds to generate various breathing biomarkers, including breathing phase, rate, depth, and symmetry [31]. WiRelax stands out as the first non-contact respiratory biofeedback system, which utilizes WiFi to map changes in Channel State Information to the instantaneous breathing state [19]. Tanaka's research focuses on real-time breathing training for individuals with mild cognitive impairment, using a tablet's camera to analyze face blood flow [42]. Additionally, the Breathing-Mentor app combines effective visualization instructions with biofeedback on deep abdominal breathing, leveraging the accelerometers in mobile phones [11]. Furthermore, Prana company utilizes a strap sensor as a wearable platform for training various at-rest breathing exercises, providing live sensor feedback and through a gamified mobile app [30].

Despite the significant progress made in real-time feedback systems for breathing control, none of the mentioned works have utilized the powerful Photoplethysmography (PPG) sensor. The PPG sensor has the capability to record interactions between cardiac, respiratory, and autonomic functions [2]. Numerous research studies have confirmed that PPG signals encompass distinct mod-

ulations, including respiratory induced intensity variations (RIIV), respiratory induced amplitude variations (RIAV), and respiratory induced frequency variations (RIFV) [7,17]. Moreover, PPG sensors integrated into wearable devices, like smartwatches, offer convenient and continuous real-time breathing monitoring. With no additional hassle, users can wear these devices and access their respiratory data seamlessly. This portability enables monitoring during various activities, making them beneficial for athletes, swimmers, singers, and anyone interested in tracking their breathing patterns for health and performance improvement. Additionally, PPG sensors offer convenience due to their non-invasive nature, allowing for comfortable and hassle-free measurements. This combination of accessibility, ease of use, and continuous monitoring makes PPG-based systems a valuable tool for various applications.

Many previous applications incorporating the PPG sensor have primarily focused on detecting parameters such as blood oxygen saturation [4,22,34,39], heart rate [22,34], blood pressure [18,22], and respiration rate [17,25,36,39]. Among these studies, Karlen [17] employed a smart fusion technique to estimate RR using high-sampling rate PPG with 1-second window shifts, achieving an average root mean square error (RMSE) of 3 bpm. Park [25] utilized an adaptive lattice notch filter and high-sampling rate PPG to estimate RR, providing updated results every 5 seconds. In a more recent study by Selvakumar [36], an incremental merge segmentation technique with high sampling rate PPG and 5-second shifting windows resulted in Mean Absolute Error (MAE) of less than 1 bpm, outperforming previous methods. While the aforementioned studies primarily focused on real-time respiratory rate estimation, none of them specifically addressed the tracking of the breathing cycle as a significant variable. Notably, a recent review by Alian [3] introduced the clinical usage of PPG but did not address the monitoring of the breathing cycle. This underscores the existing research gap and highlights the significance of developing innovative approaches that leverage the potential of the PPG sensor to enable comprehensive and real-time monitoring of the breathing cycle.

This study aims to develop and validate a smart, low-power wearable device utilizing reflectance-type PPG that can monitor the breathing process in real-time. Recently, there has been a growing trend towards the utilization of machine learning techniques in real-time processing of biophysiological data [9,32,40,43], despite the prevalence of signal processing methods in previous studies. Machine learning methods offer significant advantages in handling and analyzing biophysiological data in real-time scenarios. In this study, various signal processing and machine learning algorithms were evaluated to extract respiratory information from raw PPG signals. Among these methods, the Empirical Mode Decomposition (EMD) was tested as a baseline. Additionally, a hybrid model combining an autoencoder (AE) and sequence-to-sequence (Seq2Seq) models was explored. Sequence-to-sequence models were considered due to the input and ground-truth signals being from different domains. Furthermore, a conditional Generative Adversarial Network (GAN) with Wasserstein loss was employed to generate real-time breathing curves from PPG signals. The results revealed that

the GAN network outperformed the other two methods, showing its potential as an effective and efficient approach for respiratory monitoring using PPG signals. This GAN model has been specifically designed to effectively operate on PPG devices with very low sampling rates, making it adaptable to other PPG devices. The generated breathing signal can be utilized to extract valuable information regarding the breathing process, including parameters such as breathing depth, inhalation and exhalation durations, or any other relevant data based on the specific application requirements. In addition to this, to further enhance the monitoring of breathing, a peak detection algorithm was employed to estimate RR and breathing depth after generating the breathing curve. By employing this approach, the proposed wearable device offers a cost-effective solution for real-time monitoring of the breathing cycle and RR in individuals.

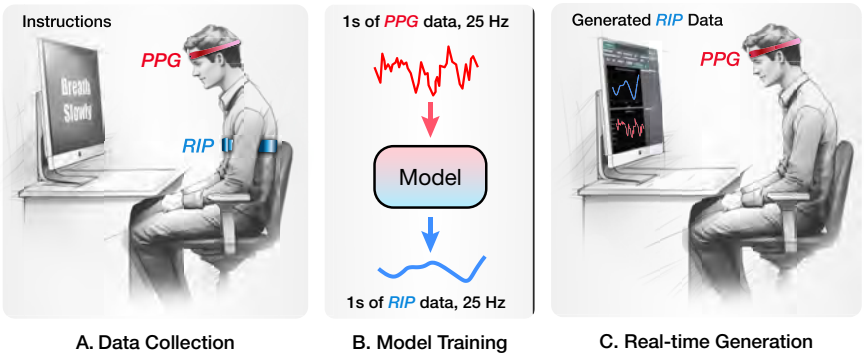

Fig. 1. Overview of the system: A) In data collection, participants are instructed to breathe normally while data from the PPG sensor and Respiratory Inductive Plethysmography (RIP) sensor are recorded. B) Our algorithms, namely Empirical Mode Decomposition (EMD), Sequence-to-Sequence (Seq2Seq), and Generative Adversarial Network (GAN), will be trained utilizing 1-second windows of pre-processed PPG data and RIP signal as input and ground truth, respectively. C) In inference time, respiratory cycle is tracked only using PPG signal and subjects can actively monitor their breathing cycle and respiratory rate (RR) values every 1 second after 20 seconds of data acquisition.

2 Methods

2.1 Overview

Real-time feedback on breathing phases and respiratory rate holds significant potential for clinical, at-home and general ambulatory applications. However, accurately tracking respiration using a comfortable, low-cost, low-energy platform may pose challenges compared to the precision achieved by devices em-

ployed in hospital settings. In this study, we propose an algorithm that leverages a comfortable headband embedded with low-frequency PPG technology.

Fig. 1 provides an overview of our system. In the data collection phase, participants are instructed to breathe normally while data from the PPG sensor in the headband and accelerometer data from the chest strap, Respiratory Inductive Plethysmography (RIP) sensor are recorded (Fig. 1.A). The 1s windows of pre-processed PPG data and RIP signal are used as input and ground truth, respectively (Fig. 1.B), for the proposed algorithms namely Empirical Mode Decomposition (EMD), Sequence-to-Sequence (Seq2Seq), and Generative Adversarial Network (GAN). The performance of each algorithm in accurately tracking breathing is evaluated, and the output from the most effective algorithm is utilized by the respiration-rate estimator.

During the feedback phase using the trained models (Fig. 1.C), subjects can monitor their breathing cycle with a delay of 1 second. Respiratory rate (RR) values are generated after 20 seconds of data acquisition every 1s, providing subjects with timely feedback on their respiration.

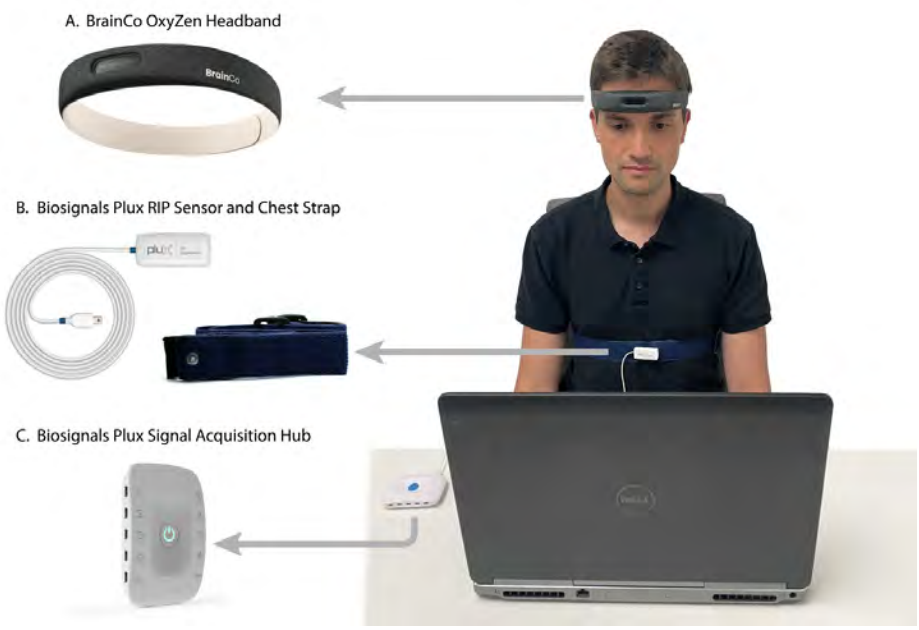

Fig. 2. Experimental setup: Subjects were instructed to breath normally while synchronized data of A) PPG sensor embedded in BrainCo's OxyZen headband and B) Biosignals Plux RIP sensor in conjunction with C) a signal acquisition hub were collected.

2.2 Experimental Setup

For the purpose of this preliminary study, 10 subjects participated (3 females, age 27.4 ± 4.83 years). Participants were instructed to assume a comfortable seated position and breathe normally for 9 minutes. Fig. 2, shows an example of the experimental setup. Subjects participating in the study were instructed to wear a BrainCo's OxyZen headband [8,12], which incorporated a PPG sensor that uses 0.74mW/s with 25 Hz sampling rate. Additionally, they were instructed to wear a Biosignals Plux RIP sensor [29] with 1000 Hz sampling rate. RIP is a non-invasive technique used to measure respiratory function. The output of the RIP sensor is a continuous waveform that represents the changes in thoracoabdominal dimensions over time, providing valuable information about a person's breathing behavior. All recorded data from PPG and RIP sensor was then transferred to the laptop using bluetooth and all processing was performed using python programming language.

2.3 Pre-processing

We conducted data pre-processing steps to prepare the RIP and PPG data for analysis. The RIP signal, originally acquired at a sampling rate of 1000 Hz, was downsampled to match the sampling rate of the head-worn PPG sensor, which operated at 25 Hz. In order to prepare the inputs and outputs of the models we used a window-based approach, generating 1s (25 samples) windows of input and output with 20 samples of overlap between hops. Then, we applied a min-max normalization to each 1 second signal of respiratory strap and PPG data.

2.4 Respiratory Signal Generation

EMD. Empirical Mode Decomposition (EMD) is a powerful data analysis method that decomposes a signal into intrinsic mode functions (IMFs) through an iterative sifting process. Each IMF represents a component of the signal with a distinct frequency and timescale [45]. We employed Ensemble EMD (EEMD) which is an improved version of EMD and is previously used in respiratory rate estimation [23,37]. EMD can be sensitive to noise and initial conditions, which may result in variations and inconsistencies in the decomposed IMFs. Ensemble EMD addresses these limitations by introducing an ensemble approach. It involves performing multiple iterations of the EMD process on the same signal with different noise realizations or initial conditions. By averaging or combining the resulting IMFs across iterations, Ensemble EMD aims to improve the overall decomposition quality, reduce artifacts, and enhance the reliability of the extracted components [37]. The key idea behind Ensemble EMD is to leverage the variations introduced by different noise realizations or initial conditions to enhance the robustness of the decomposition.

The use of EEMD has been explored extensively in post-processing respiratory rate calculations [16]. Therefore we assess EEMD as a strong baseline of comparison. EEMD was applied to the PPG signal to extract the IMFs using

publicly available PyEMD library [27,28]. To reduce processing time and accommodate real-time constraints, we imposed a limitation of 5 iterations during the EMD process. Additionally, our data windows were restricted to 1 second to ensure real-time feasibility. As a result, we generated only 2 IMFs during the EMD. Based on visual inspection and comparison, we identified the second IMF, which exhibited the closest resemblance to the respiratory curve and contained the most relevant respiratory information, as the chosen IMF for further analysis and tracking.

Seq2Seq. After a series of preliminary experiments using simple feedforward, recurrent and convolutional architectures, we developed a hybrid model combining an autoencoder and a Seq2Seq model. An autoencoder (AE) is a type of unsupervised artificial neural network that aims to reconstruct its input at the output layer while imposing a condition on its hidden representations (e.g. size, distribution, sparsity). Autoencoders consist of an encoder that compresses the input data into a lower-dimensional representation, called the latent vector, while the decoder reconstructs the input from this latent representation. The use of autoencoders offers several advantages, such as dimensionality reduction, feature extraction, and new data generation. By compressing the input data into a lower-dimensional space, autoencoders can capture the most salient features and eliminate noise and redundant information [5,6].

Sequence-to-sequence is a type of supervised artificial neural network that aims to transform an input sequence to an output sequence. The Seq2Seq model also consists of an encoder and a decoder. The encoder processes the input sequence and converts it into a fixed-length latent representation [13,41].

In this study we used a Seq2Seq model in combination with an AE. Due to the different nature of the input and output sequences (ground-truth being a smooth signal and PPG containing fast oscillations), using only the Seq2Seq model resulted in high losses. To address this, an AE was introduced to extract the most important information or latent vector from the ground-truth signal. By generating a more sparse representation of the output signal, the AE helps to better match the characteristics of the PPG signal, making the overall model more effective in generating accurate results and reducing the losses encountered in the Seq2Seq approach. In this approach, the autoencoder accepts inductive respiratory chest strap signals as input and encodes the latent vector representation of the RIP signals. The latent representation becomes the ground truth for our Seq2Seq model, which utilizes the raw PPG to predict the encoded latent vector from the autoencoder to generate respiratory information from a PPG signal. The Seq2Seq structure then outputs a generated latent vector, which is passed as input to the previously trained decoder of our autoencoder to create generated respiratory signals.

Fig. 3 represents the schematic of Seq2Seq model. The Seq2Seq component (Fig. 3.A) starts with an input layer, followed by 2 consecutive Conv1D blocks. An add layer is applied to combine the outputs of two convolutional layers. This is followed by additional Conv1D layers with layer normalization and add layers

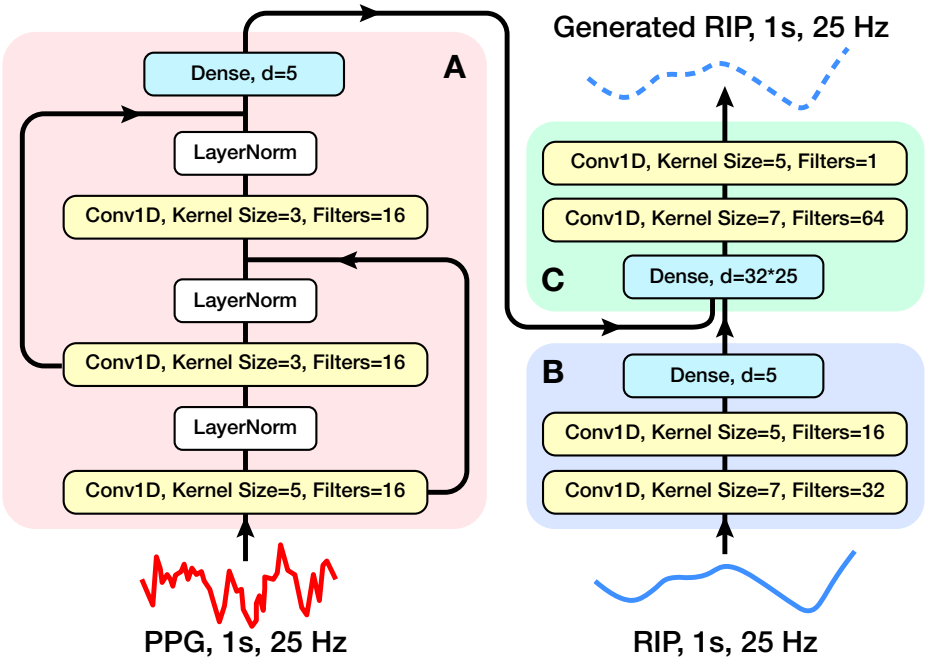

Fig. 3. The Seq2Seq model incorporates an autoencoder (AE) for RIP signals and an encoder for PPG, resulting in a hybrid model structure. A) PPG encoder: encodes raw PPG into latent representation of respiratory signals. B) RIP endoder: encodes RIP signals into latent representations for the decoder. This block was only used during training and was discarded at inference time C) RIP decoder: The generated latent representation of respiratory signals is given to the decoder to generate the respiratory signal. At inference time, the input signal to the decoder comes from the PPG encoder and the RIP encoder is discarded.

to further refine the representations. The output is then flattened and passed through a dense layer for final processing.

The autoencoder begins with an input layer, followed by a sequential layer. The encoder section (Fig. 3.B) of the autoencoder includes Conv1D layers and a dense layer for dimensionality reduction. The encoder is followed by a sequential layer then the decoder (Fig. 3.C) which is composed of dense layer and Conv1D layers for signal reconstruction. The whole model was developed using TensorFlow [1].

This approach enabled us to generate respiratory signals that closely resembled the original data. We utilized an Adam optimizer with a learning rate of 0.0001 and gradient clipping with a threshold of 0.5 to stabilize training. The mean squared error (MSE) loss function was employed to measure the discrepancy between the predicted respiratory signals and the ground truth. This helped ensure that the model was trained sufficiently without compromising its generalization capabilities.

GAN. A Generative Adversarial Network (GAN) is a type of machine learning model that consists of two main components: a generator and a discriminator. The generator aims to synthesize realistic samples, while the discriminator learns to distinguish between real and generated samples [14]. By training the generator to produce respiratory signals that closely resemble real signals, GANs offer a powerful tool for simulating realistic respiratory patterns. In this process, both parts of the GAN (generator and discriminator) are used during training. However, during the testing phase, the discriminator is no longer utilized for generating the output, resulting in significantly faster processing times. This efficiency allows the GAN-based model to generate accurate and realistic respiratory information in real-time applications. A Conditional GAN introduces additional conditioning information to both the generator and the discriminator, allowing for more targeted and controlled generation of samples [21].

A Conditional GAN will take pairs of real data samples and their conditions as well as pairs of fake or generated samples and intended conditions for training. Using feedback from the discriminator, the generator learns to produce data that is more realistic and aligns well with a given data pair condition. In our model, the positive pair (pertaining to ground truth samples) was defined as real PPG and RIP sensor signals. The negative pair (pertaining to generated samples) was defined as real PPG and generated RIP sensor signals.

In GANs, loss functions quantify the discrepancy between the generated samples and the real samples, providing a signal for the model to optimize its performance. Two commonly used loss functions in GANs are cross-entropy loss and Wasserstein loss. Cross-entropy loss measures the dissimilarity between the predicted distribution and the true distribution [46]. Wasserstein loss calculates the distance between the generated and real distributions using the Earth Mover's Distance. To stabilize training and encourage the generator to produce diverse samples, gradient penalty is utilized in conjunction with Wasserstein loss GANs. Gradient penalty introduces a regularization term that penalizes the discrimi-

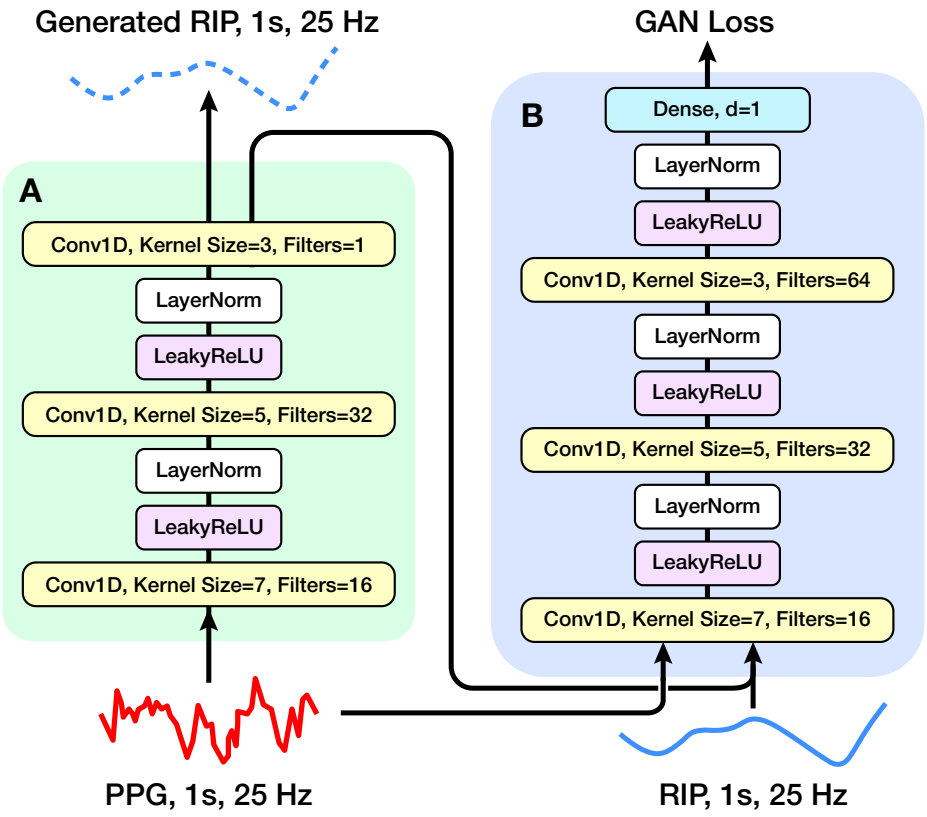

Fig. 4. Conditional GAN model showing a generator block and a discriminator block, A) The Generator tries to generate RIP signals based on the PPG conditioning input. B) the Discriminator computes the discriminator loss based on the input conditioning signal and the ground-truth RIP signal. If the input to the discriminator comes from a pair of ground truth RIP and its corresponding PPG, the block should detect the input as real; otherwise, if the generated RIP is paired with PPG, the discriminator should classify it as fake.

nator for having large gradients, preventing it from overpowering the generator during training [15]. We employed Wasserstein loss, which proved to be effective in mitigating mode collapse and promoting diverse sample generation. We further incorporated gradient penalty to enhance generated signal quality. We used an RMSprop optimizer with a learning rate of 5e-5. The model was trained for 200 epochs.

In our model depicted in Fig. 4, the generator component (Fig. 4.A) begins with an input layer, followed by a series of Conv1D layers. Each Conv1D layer is paired with a Leaky ReLU nonlinearity and LayerNorm. The discriminator component (Fig. 4.B) also starts with an input layer, followed by multiple convolutional layers with a Leaky ReLU nonlinearity and LayerNorm layers. This model was developed using TensorFlow [1].

2.5 Respiratory Rate Estimation

To confirm the accuracy of the generated signal, an offline analysis of respiratory rate (RR) estimation on the 1 minute test signal was performed. To accurately estimate RR from the real and generated signals obtained through our proposed system, we employed a custom peak detection pipeline. Firstly, we applied detrending to the signals, which helped improve the definition of peak occurrences, enhancing the accuracy of peak detection. Next, a moving average with a 1-second window was applied to the signals. This moving average served to smoothen the data and reduce any noise or fluctuations, providing a cleaner signal for further analysis. We then identify peak occurrences across the signals and store their indices to an array. To detect the valleys between the peaks, we determine the signal values between two consecutive peak indices and return the index of the minimum value. This algorithm effectively identified the peaks present in the signals, which corresponded to the peaks in the breathing cycles.

Once the peaks and valleys were identified, we proceeded to calculate the respiratory rate. This involved determining the number of peaks occurring within a given window range defined as peaks per minute. To achieve this, we counted the number of peaks within a specified window length of the signal and multiplied this rate by the sample rate to obtain the rate in peaks per second. Finally, we multiplied this value by 60 seconds to express the respiratory rate as peaks per minute. We calculated respiratory rate in windows of 30-seconds with a 29 second overlap. We defined the values to represent the breaths per-minute respiratory rate.

2.6 Breathing Depth

Breathing depth plays a significant role in differentiating between deep and shallow breathing. It serves as an indicator of the amount of inhaled air that fills the lungs during respiration [31]. To quantify breathing depths, we employed an amplitude measurement, which represents the difference between the peak and valley of each breath. For each subject, we calculated the breathing depth for every breath, obtaining individual depth scores. To compare the ground truth

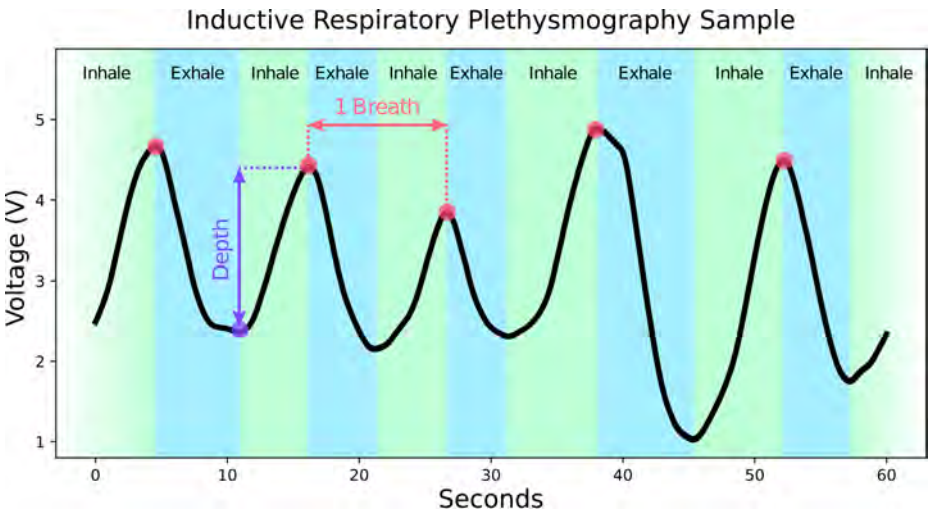

Fig. 5. Breathing curve characteristics: a 1-minute sample of the RIP signal is depicted with the peaks highlighted. Here RR, defined as the number of peaks in a minute, is five bpm. First, the distance between consecutive peaks and valleys is calculated for the depth score, and then this difference is averaged over the entire sequence.

breathing depths with the generated breathing depths, we computed the average breathing depth scores for each subject. This allowed us to assess the accuracy and consistency of the generated respiratory signals in capturing the desired breathing depths.

Fig. 5 demonstrates an example of breathing curve, highlighting the signal peaks and what is denoted as 1 breath and breathing depth.

2.7 Real-Time Implementation

The generated respiratory signal was streamed at a rate of 25 Hz. To enhance the accuracy of the signal and reduce high-frequency noise, we applied a 1 second window moving average. Smoothing the generated signal results in a more reliable representation of the respiratory pattern. Fig. 6 illustrates an example of the streamed generated signal (in blue) along with the recorded PPG signal (in red). We then apply the aforementioned peak detection algorithm to identify respiratory rate as using a 20-second window length updated every 1 second. Next, we performed a weighted classification of the breathing phase based on the signal characteristics. By analyzing the slope of the signal, we classified it as either breathing in or breathing out. A positive slope indicated the onset of inhalation, while a negative slope indicated the onset of exhalation. The duration of the positive or negative slope influenced the weighting or confidence assigned to each breathing phase classification. A longer duration of a particular slope resulted in a higher weighted classification, indicating a higher level of confidence

in the assigned phase. Breathing phase was detected every 1 second and stored in a phase classification array.

Fig. 6. GUI: In the study focusing on the breathing generation based on PPG, the PPG- BREATHING tab of the developed SDK is used. Other tabs won't be discussed as they are out of the scope of this paper. In this tab, the model-generated respiratory curve is displayed at top and Red PPG curve at the bottom. Respiratory Rate estimated based on our generated signal is shown beside the top plot.

To identify a complete breath, we analyzed the stream of breathing state classifiers. By monitoring the transitions between breathing in and breathing out states, we determined the occurrence of a complete breath. This sliding 20-second window (with an overlap of 19-seconds) allowed for a dynamic classification of breathing phase.

3 Results

In this study, we utilized three different methods to track the breathing cycle. Fig. 7 shows an example of the breathing signals generated by these methods (Fig. 7.C), along with PPG (Fig. 7.A) and the ground truth signal (Fig. 7.B) recorded from a RIP sensor. The models produced real-time signals every second, resulting in 60 windows that were then concatenated, smoothed and detrended.

Since PPG and RIP have different ranges, for visualization all generated signals were min-max normalized.

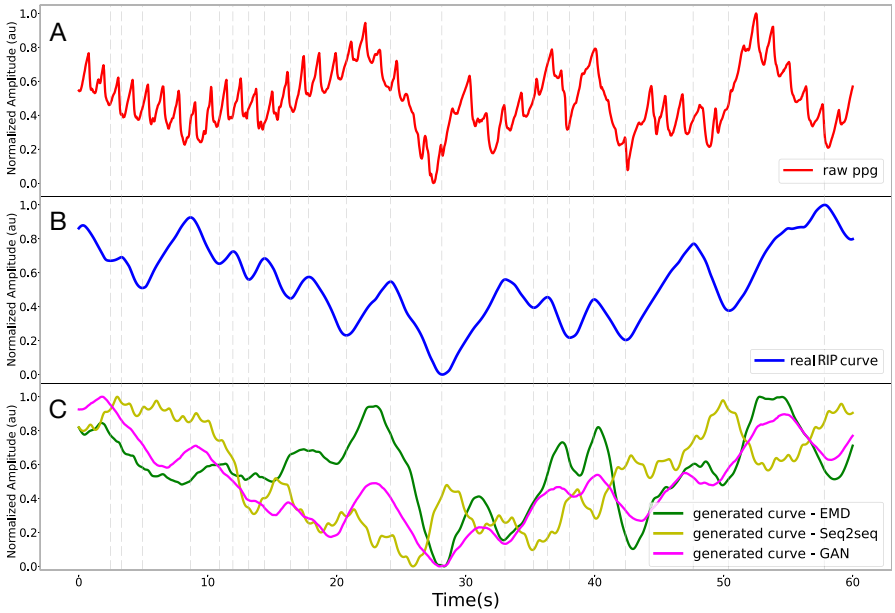

Fig. 7. Model performance comparison: A) A sample of PPG signal given as the input to the models, B) The ground-truth signal with vertical dashed lines highlighting the start and end of each breathing phase (inhale or exhale). C) Generated signals by the three models for the given PPG signal. The GAN-generated signal is more accurately following the ground-truth.

It is important to note that the generated signals have a 1-second delay compared to the real signal. The example in the figure was intentionally selected to demonstrate the strengths and limitations of our proposed method. Since the EMD method does not rely on supervised learning, its output resembles the PPG signal more than the desired RIP signal. As shown in the figure, the GAN-generated signal closely follows the ground truth compared to the other two methods. However, the GAN algorithm faces challenges in tracking very shallow breaths (<0.01 au) and fast breaths with higher frequencies (>20 bpm). Since any breathing period with a bpm above 20 is defined as non-normal breathing, rapid breathing, or tachypnea [26] we define this algorithm as qualified for the normal breathing range.

To quantitatively assess the performance of three methods in tracking the ground-truth breathing cycle, Dynamic Time Warping (DTW) and Mean Absolute Error (MAE) were utilized. DTW is a technique employed to measure the

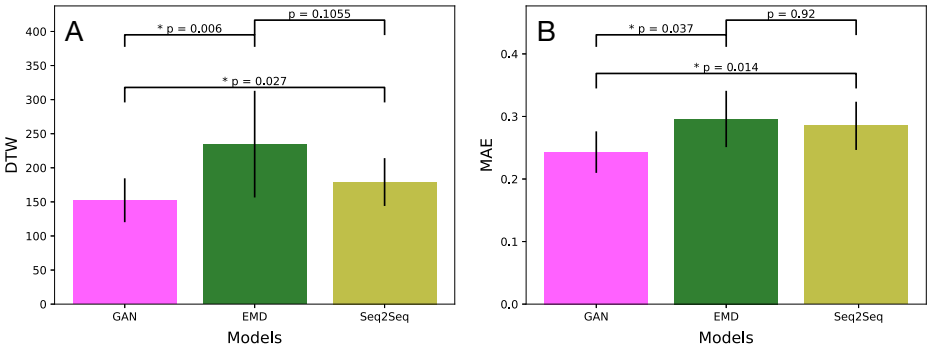

Fig. 8. Quantitative comparison of models: A) As one measure of dissimilarity between generated signal and ground-truth signal, average DTW distances over 1-minute test signal for all subjects were compared for all three methods, GAN method shows better performance with statistical significance. B) All three method's MAE between generated and ground-truth signals were also compared to reflect the error between the two. Similarly GAN-generated signal with less error showed closer behavior to ground-truth with statistical significance.

similarity between two time series sequences of different lengths. In comparison to cross-correlation, DTW excels in local pattern matching, enabling it to focus on identifying similar segments of the signals. This effectiveness becomes evident in scenarios where certain segments hold greater importance than others [33]. Specifically, in our generated data, where the signal closely resembles the ground-truth during specific segments of normal breathing, DTW distance served as one of the measures to evaluate the performance of the methods quantitatively. Generated respiratory signals using each method were compared with ground-truth over the 60s test data. As a second measure, the normalized generated signals were compared to the ground truth using MAE. MAE is a metric that calculates the average absolute difference between the generated signals and the ground truth, providing a quantitative measure of the overall accuracy of the generated signals in relation to the ground truth. In order to compare the statistical difference, the results were then tested for normality using the Shapiro-Wilk test showing non-normal behavior in some distributions. To simultaneously compare the performance of the three methods, a Friedman rank test is utilized, followed by a post hoc pairwise test employing the Wilcoxon signed-rank test. Lastly, Holm's method is used to adjust the significance thresholds (alpha). All statistical analysis were performed using SciPy [44] and statsmodels [35] libraries. The results of DTW and MAE analysis are shown in Fig. 8 for test data. Both measures reveal that the GAN performance surpasses the other two methods with statistical significance.

After generating the breathing curves, we proceeded to estimate the RR for each subject using all methods. In Fig. 9.A, we compared the average RR estimation over a 30-second sliding window of the generated test signals with the ground-truth signal for each subject. Furthermore, we calculated the MAE be-

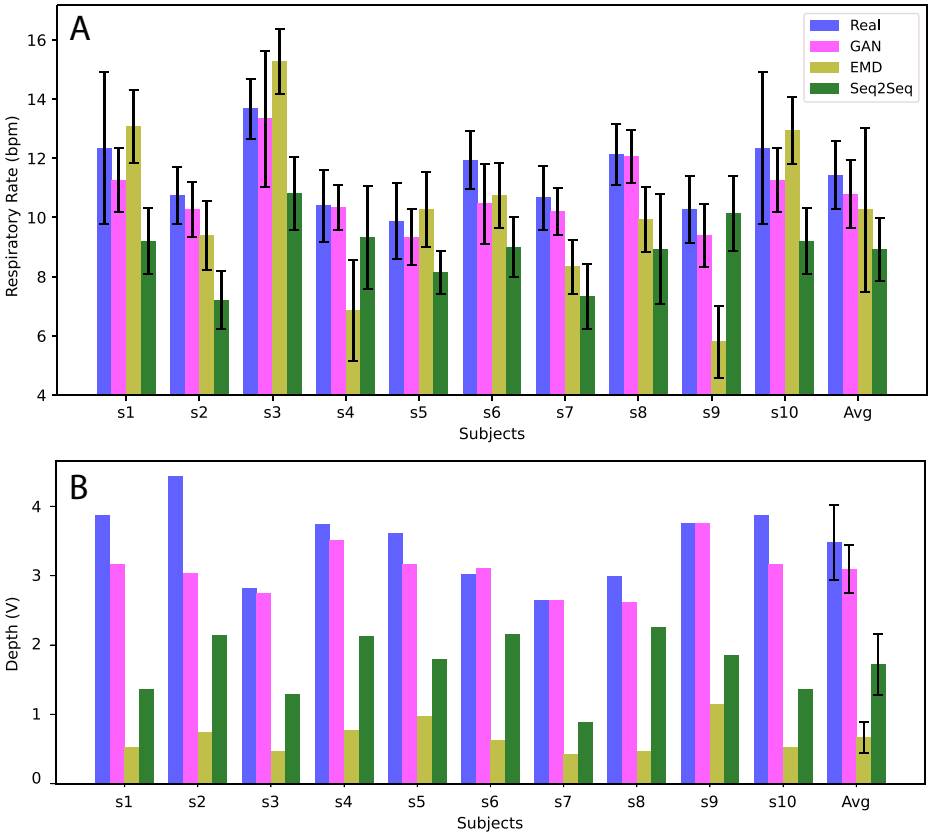

Fig. 9. Breathing measures: A) Average RR estimation over 1 minute test signal, based on the generated signals by three models are compared with RR of ground-truth for all subjects individually as well as average over all subjects. B) Depth scores for generated signals and the ground-truth curve are compared for all individual subjects and average of all subjects. These graphs shows the superior performance of GAN-generated signal in monitoring breathing pattern.

tween repetitions for each subject. This metric provides an assessment of the consistency and accuracy of the RR estimates across different repetitions. The average MAE overall subjects for GAN, EMD and Seq2Seq is 1.47, 2.28 and 2.83 bpm respectively. The MAE values provide insights into the overall variability and precision of the RR estimations, contributing to the assessment of the reliability of the GAN-generated signals in tracking and estimating respiratory rate. The results demonstrate minimal differences between the GAN-generated and ground-truth signal, indicating a high level of similarity in the RR estimations. Alongside the estimation of respiratory rate, we also computed the average breathing depth during the one-minute test data for each subject. Fig. 9.B displays the corresponding results for each subject, comparing the values obtained from the real RIP signal with those generated by our algorithm. These results underscore the capability of the GAN generated signals to closely track and mimic the actual breathing depth. The two other methods show inconsistent behaviour considering different features. EMD is closer to real signal in RR estimation compared with Seq2Seq, while Seq2Seq is better in estimating breath depths compared with EMD.

To evaluate the real-time usage of our method, we recorded demos from two subjects, one male and one female, while they were actively breathing. These demos provide valuable insights into how our method performs in real-time scenarios and can be used to showcase its potential applications and benefits. The demo video is available at this hyperlink.

4 Discussion

Real-time monitoring of breathing is of significant importance, particularly in outpatient settings where low-cost wearables that are comfortable and easy to wear are essential. In this study, we developed an algorithm utilizing a low-frequency PPG sensor, enabling real-time assessment of breathing with a 1-second latency.

Our results demonstrate the feasibility of utilizing a low-frequency PPG for monitoring breathing. The algorithm's performance was evaluated on 10 subjects, revealing the GAN's ability to generate a reliable tracking of the real respiratory signal despite the low sampling rate of the PPG (25 Hz), which resulted in a low number of samples per window (25 samples). Importantly, the GAN outperformed the other two methods, namely EMD and Seq2Seq. This is attributed to the specific challenge of the low-frequency nature of the PPG signal (25 Hz). While EMD and Seq2Seq focused on capturing the low-frequency oscillations present in the PPG, only the GAN was capable of effectively tracking the relatively higher frequency content present in the ground-truth signal. The dissimilarity between the three methods and the ground-truth signal was quantified using DTW and MAE, acknowledging that some level of variation is expected due to the different sources of the signals (RIP vs. PPG). Although the magnitude of the DTW distance itself does not determine the method's goodness, it is evident from Fig. 8.A that the GAN exhibits less dissimilarity and statistically

significant differences compared to the other methods. Fig. 8.B further supports these observations by comparing the performance of the three methods using MAE. Fig. 9.A illustrates the close resemblance of the GAN-based RR estimations to the ground truth, with an average mean absolute error (MAE) of 1.47 bpm. Additionally, Fig. 9.B demonstrates the accuracy of the GAN-generated signals in estimating breathing depth.

Table 1. Comparison of our method and Selvakumar[36]

Paper	RR MAE	Sampling Rate	Window Size
Karman et al.	1.47 bpm	25 Hz	1s
Selvakumar et al.	1 bpm	125 Hz	5s

In the recent literature, there is a scarcity of studies using low-frequency PPG for real-time breathing monitoring. In comparison to Selvakumar [36], their method achieved MAE less than 1 bpm over various breathing frequency ranges. However, it should be noted that their approach employed a PPG with a higher sampling rate of 125 Hz and relied on 5 seconds of data (625 samples per window), whereas our algorithm achieves comparable accuracy with only 1 second of data using a 25 Hz PPG sensor (25 samples per window). Overall, our results align with the latest research in the field and are applicable to any PPG-based system, given the 25 Hz sampling rate.

One of the limitations of our method is that it may encounter challenges in generating the respiratory signal accurately when the subject exhibits fast and shallow breathing, as can occur during talking or abnormal breathing patterns. One possible approach to addressing this limitation would involve utilizing a higher sampling rate for the PPG data. With a higher sampling rate, our method may have better captured the nuances of fast and shallow breathing patterns, thereby potentially improving the generation of the respiratory signal in such scenarios. Additionally, conducting a real-time evaluation study would allow us to assess the algorithm's performance in real-world scenarios. Moreover, it is important to investigate the user experience and ease of use for operators. This setup and real-time algorithm will be employed in a clinical study conducted in a hospital to aid children with pain control after surgery, with the aim of reducing the use of opioids.

In this study, we explored different signal processing and machine learning algorithms, such as EMD, Seq2Seq, and GAN, to track the breathing process in real-time using a low-power PPG signal mounted on a headband. Our developed GAN algorithm showed promising results in accurately tracking the breathing cycle. Our findings demonstrate the effectiveness of our algorithm for real-time and low-latency breathing monitoring using low-frequency PPG, highlighting its potential applications in pain control and mindful breathing practices. PPG sensors integrated into wearable devices, such as smartwatches, provide a con-

venient and continuous means for real-time breathing monitoring. Leveraging this technology presents an opportunity to track the breathing cycle in diverse situations, prompting the need for user-friendly solutions. While inductive chest strap sensors are commonly used in hospital settings, their discomfort limits their widespread adoption. Our proposed tool addresses these limitations by offering an easy-to-deploy, comfortable, and user-friendly wearable device for real-time breathing monitoring. It provides an effective and accessible solution for at-home respiratory monitoring with potential applications in various healthcare domains. This sets the stage for future research conducting online evaluation studies to further validate and improve our method.

References

1. Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., Zheng, X.: TensorFlow: Large-scale machine learning on heterogeneous systems (2015), `https://www.tensorflow.org/`, software available from tensorflow.org
2. Alian, A.A., Shelley, K.H.: Photoplethysmography. Best Practice Research Clinical Anaesthesiology **28**, 395–406 (12 2014). https://doi.org/10.1016/j.bpa.2014.08.006
3. Alian, A.A., Shelley, K.H.: Ppg in clinical monitoring. In: Photoplethysmography, pp. 341–359. Elsevier (2022)
4. Azmal, G.M., Al-Jumaily, A., Al-Jaafreh, M.: Continuous measurement of oxygen saturation level using photoplethysmography signal (2006). https://doi.org/10.1109/ICBPE.2006.348646
5. Baldi, P.: Autoencoders, unsupervised learning, and deep architectures. ICML Unsupervised and Transfer Learning (2012). https://doi.org/10.1561/2200000006
6. Barot, V., Patel, D.R.: A physiological signal compression approach using optimized spindle convolutional auto-encoder in mhealth applications. Biomedical Signal Processing and Control **73**, 103436 (3 2022). https://doi.org/10.1016/j.bspc.2021.103436
7. Boccignone, G., D'Amelio, A., Ghezzi, O., Grossi, G., Lanzarotti, R.: An evaluation of non-contact photoplethysmography-based methods for remote respiratory rate estimation. Sensors **23**, 3387 (3 2023). https://doi.org/10.3390/s23073387
8. BrainCo: `https://brainco.tech/`, last accessed 24 July 2023
9. Chen, L., Liu, X., Peng, L., Wu, M.: Deep learning based multimodal complex human activity recognition using wearable devices. Applied Intelligence **51** (2021). https://doi.org/10.1007/s10489-020-02005-7
10. Drigas, A., Mitsea, E.: Breathing: A powerfull tool for physical & neuropsychological regulation. the role of mobile apps. Technium Soc. Sci. J. **28**, 135 (2022)
11. Faust-Christmann, C.A., Taetz, B., Zolynski, G., Zimmermann, T., Bleser, G.: A biofeedback app to instruct abdominal breathing (breathing-mentor): pilot experiment. JMIR mHealth and uHealth **7**(9), e13703 (2019)
12. FocusCalm: `https://focuscalm.com/`, last accessed 24 July 2023
13. Gehring, J., Auli, M., Grangier, D., Yarats, D., Dauphin, Y.N.: Convolutional sequence to sequence learning (5 2017)

14. Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial networks (6 2014)
15. Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., Courville, A.: Improved training of wasserstein gans (3 2017)
16. Hadiyoso, S., Dewi, E.M., Wijayanto, I.: Comparison of emd, vmd and eemd methods in respiration wave extraction based on ppg waves. Journal of Physics: Conference Series **1577**, 012040 (7 2020). https://doi.org/10.1088/1742-6596/1577/1/012040
17. Karlen, W., Raman, S., Ansermino, J.M., Dumont, G.A.: Multiparameter respiratory rate estimation from the photoplethysmogram. IEEE Transactions on Biomedical Engineering **60**, 1946–1953 (7 2013). https://doi.org/10.1109/TBME.2013.2246160
18. Khalid, S.G., Zhang, J., Chen, F., Zheng, D.: Blood pressure estimation using photoplethysmography only: Comparison between different machine learning approaches. Journal of Healthcare Engineering **2018**, 1–13 (10 2018). https://doi.org/10.1155/2018/1548647
19. Khamis, A., Kusy, B., Chou, C.T., Hu, W.: Wirelax: Towards real-time respiratory biofeedback during meditation using wifi. Ad Hoc Networks **107**, 102226 (10 2020). https://doi.org/10.1016/j.adhoc.2020.102226
20. Lukic, Y.X., Teepe, G.W., Fleisch, E., Kowatsch, T.: Breathing as an input modality in a gameful breathing training app (breeze 2): Development and evaluation study. JMIR Serious Games **10**, e39186 (8 2022). https://doi.org/10.2196/39186
21. Mirza, M., Osindero, S.: Conditional generative adversarial nets (11 2014)
22. Mohan, P.M., Nisha, A.A., Nagarajan, V., Jothi, E.S.J.: Measurement of arterial oxygen saturation (spo¡inf¿2¡/inf¿) using ppg optical sensor. pp. 1136–1140. IEEE (4 2016). https://doi.org/10.1109/ICCSP.2016.7754330
23. Motin, M.A., Karmakar, C.K., Palaniswami, M.: Ensemble empirical mode decomposition with principal component analysis: A novel approach for extracting respiratory rate and heart rate from photoplethysmographic signal. IEEE Journal of Biomedical and Health Informatics **22**, 766–774 (5 2018). https://doi.org/10.1109/JBHI.2017.2679108
24. Nicolò, A., Massaroni, C., Schena, E., Sacchetti, M.: The importance of respiratory rate monitoring: From healthcare to sport and exercise. Sensors **20**(21), 6396 (2020)
25. Park, C., Lee, B.: Real-time estimation of respiratory rate from a photoplethysmogram using an adaptive lattice notch filter. BioMedical Engineering OnLine **13**, 170 (2014). https://doi.org/10.1186/1475-925X-13-170
26. Park, S.B., Khattar, D.: Tachypnea (2023)
27. Pele, O., Werman, M.: A linear time histogram metric for improved sift matching. In: Computer Vision–ECCV 2008. pp. 495–508. Springer (October 2008)
28. Pele, O., Werman, M.: Fast and robust earth mover's distances. In: 2009 IEEE 12th International Conference on Computer Vision. pp. 460–467. IEEE (September 2009)
29. Plux: `https://www.pluxbiosignals.com/products/inductive-respiration-rip-sensor`, last accessed 24 July 2023
30. Prana.co: `https://prana.co/`, last accessed 24 July 2023
31. Rahman, M.M., Ahmed, T., Ahmed, M.Y., Dinh, M., Nemati, E., Kuang, J., Gao, J.A.: Breathebuddy: Tracking real-time breathing exercises for automated biofeedback using commodity earbuds. Proceedings of the ACM on Human-Computer Interaction **6**, 1–18 (9 2022). https://doi.org/10.1145/3546748

32. Ribeiro, H.D.M., Arnold, A., Howard, J.P., Shun-Shin, M.J., Zhang, Y., Francis, D.P., Lim, P.B., Whinnett, Z., Zolgharni, M.: Ecg-based real-time arrhythmia monitoring using quantized deep neural networks: A feasibility study. Computers in Biology and Medicine **143**, 105249 (2022)
33. Sakoe, H., Chiba, S.: Dynamic programming algorithm optimization for spoken word recognition. IEEE Transactions on Acoustics, Speech, and Signal Processing **26** (1978). https://doi.org/10.1109/TASSP.1978.1163055
34. Sangeeta, B., Laxmi, S.: A real time analysis of ppg signal for measurement of spo2 and pulse rate. International Journal of Computer Applications **36** (2011)
35. Seabold, S., Perktold, J.: statsmodels: Econometric and statistical modeling with python. In: 9th Python in Science Conference (2010)
36. Selvakumar, K., Kumar, E.V., Sailesh, M., Varun, M., Allan, A., Biswajit, N., Namrata, P., Upasana, S.: Realtime ppg based respiration rate estimation for remote health monitoring applications. Biomedical Signal Processing and Control **77**, 103746 (8 2022). https://doi.org/10.1016/j.bspc.2022.103746
37. Sharma, H.: Extraction of respiratory rate from ppg using ensemble empirical mode decomposition with kalman filter. Electronics Letters **56**, 650–653 (6 2020). https://doi.org/10.1049/el.2020.0566
38. Shih, C.H.I., Tomita, N., Lukic, Y.X., Álvaro Hernández Reguera, Fleisch, E., Kowatsch, T.: Breeze: Smartphone-based acoustic real-time detection ofbreathing phases for a gamified biofeedback breathing training. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies **3**, 1–30 (12 2019). https://doi.org/10.1145/3369835
39. Shuzan, M.N.I., Chowdhury, M.H., Chowdhury, M.E.H., Murugappan, M., Bhuiyan, E.H., Ayari, M.A., Khandakar, A.: Machine learning-based respiration rate and blood oxygen saturation estimation using photoplethysmogram signals. Bioengineering **10**, 167 (1 2023). https://doi.org/10.3390/bioengineering10020167
40. Soroushmojdehi, R., Javadzadeh, S., Pedrocchi, A., Gandolla, M.: Transfer learning in hand movement intention detection based on surface electromyography signals. Frontiers in Neuroscience **16** (11 2022). https://doi.org/10.3389/fnins.2022.977328
41. Sutskever, I., Vinyals, O., Le, Q.V.: Sequence to sequence learning with neural networks (9 2014)
42. Tanaka, M., Kakuma, T., Asada, T.: Utility of paced breathing tablet guidance apparatus with real-time feedback on autonomic function for individuals with mild cognitive impairment: a pilot study. Psychogeriatrics **23**, 434–441 (5 2023). https://doi.org/10.1111/psyg.12950
43. Ullah, I., Hussain, M., ul Haq Qazi, E., Aboalsamh, H.: An automated system for epilepsy detection using eeg brain signals based on deep learning approach. Expert Systems with Applications **107**, 61–71 (10 2018). https://doi.org/10.1016/j.eswa.2018.04.021
44. Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S.J., Brett, M., Wilson, J., Millman, K.J., Mayorov, N., Nelson, A.R.J., Jones, E., Kern, R., Larson, E., Carey, C.J., Polat, İ., Feng, Y., Moore, E.W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E.A., Harris, C.R., Archibald, A.M., Ribeiro, A.H., Pedregosa, F., van Mulbregt, P., SciPy 1.0 Contributors: SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods **17**, 261–272 (2020). https://doi.org/10.1038/s41592-019-0686-2
45. WU, Z., HUANG, N.E.: Ensemble empirical mode decomposition: A noise-assisted data analysis method. Advances in Adaptive Data Analysis **01**, 1–41 (1 2009). https://doi.org/10.1142/S1793536909000047

46. Xu, J., Ren, X., Lin, J., Sun, X.: Diversity-promoting gan: A cross-entropy based generative adversarial network for diversified text generation. pp. 3940–3949. Association for Computational Linguistics (2018). https://doi.org/10.18653/v1/D18-1428

Acknowledgments. We want to acknowledge Ken Glass at Dell Technologies for providing state-of-the-art machine learning work stations.
We acknowledge Plux Biosignals for providing our data acquisition toolkit.
We also want to thank Lawrence Franchini for pushing us to tackle respiratory sensing as well as the BrainCo team members who were able to participate in our data collection procedures.