

UC Irvine

UC Irvine Electronic Theses and Dissertations

Title

Noise Removal using Deep Generative Model

Permalink

<https://escholarship.org/uc/item/34z3w988>

Author

Afandizadeh Zargari, Amir Hosein

Publication Date

2021

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

Noise Removal using Deep Generative Model

THESIS

submitted in partial satisfaction of the requirements
for the degree of

MASTER OF SCIENCE

in Computer Engineering

by

Amir Hosein Afandizadeh Zargari

Thesis Committee:
Professor Fadi Kurdahi, Chair
Associate Professor Amir Rahmani
Professor Nader Bagherzadeh

DEDICATION

I dedicate my dissertation work to my family and many friends.

A special feeling of gratitude to my loving parents, Shokouh and Shahriar whose words of encouragement and push for tenacity ring in my ears.

My sister Elham, has never left my side and is very special.

Also, this work is dedicated to my wife, Marzieh, who has been a constant source of support and encouragement during the challenges of graduate school and life. I am truly thankful for having you in my life.

TABLE OF CONTENTS

	Page
LIST OF FIGURES	iv
LIST OF TABLES	v
ACKNOWLEDGMENTS	vi
VITA	vii
ABSTRACT OF THE THESIS	viii
1 Introduction	1
2 Methods	5
2.1 BIDMC Dataset	6
2.2 Data Collection	7
2.3 Noisy PPG signal generation	8
2.4 Noise Detection	10
2.5 Noise Removal	12
3 Results	16
3.1 Comparison	18
4 Discussion	19
5 Conclusions	20
Bibliography	21

LIST OF FIGURES

	Page
2.1 Flowchart of the proposed PPG motion artifacts noise removal	5
2.2 Experimental procedure to collect accelerometer data.	7
2.3 An example of Accelerometer data from Connect, the subject moves into position, walks, runs, and then simulates the turning of a car's steering wheel. The dimensional axes are depicted in red, green and blue.	8
2.4 The same data as Figure 2.3 is visualized using the moving average. From Connect, the subject moves into position, walks, runs, and then simulates the turning of a car's steering wheel. The dimensional axes are depicted in red, green and blue.	9
2.5 The structure of the noise detection model.	10
2.6 An example of signal to image transformation.	13
3.1 The reconstructed signal of Figure 2.6 alongside with the noisy and the reference signals	17

LIST OF TABLES

	Page
2.1 The layer configuration of the noise detection model.	11
3.1 Results of the proposed method.	17
3.2 The summary comparison of our result with the existing methods. MAE stands for Mean absolute error.	18

ACKNOWLEDGMENTS

I am deeply grateful to my advisor and committee chair, Professor Fadi Kurdahi, for his invaluable advice, continuous support, and patience during my study.

I would like to express my sincere gratitude to professor Amir Rahmani, for his assistance at every stage of the research project.

I would also like to special thank Seyed Amir Hosein Aqajari and Hadi Khodabandeh, co-authors of this research, for equally contributing to this work.

I would like to extend my sincere thanks to my committee members, Professor Amir Rahmani and Professor Nader Bagherzadeh, for assigning their valuable time to review my work.

This work was partially supported by the CAREER award through ECCS-2028782 received by Fadi Kurdahi.

Finally, the original paper of this work has been submitted to Arxiv.org and ACM Health [27].Some text of this thesis is a reprint of the material as it appears in [27].

VITA

Amir Hosein Afandizadeh Zargari

EDUCATION

Master of Science in Computer Engineering University of California, Irvine	2018-2021 <i>Irvine</i>
Bachelor of Science in Computer Engineering Sharif University of Technology	2013-2018 <i>Tehran</i>

RESEARCH EXPERIENCE

Graduate Research Assistant University of California, Irvine	2018-2021 <i>Irvine</i>
Research Assistant and Data Analyst Sharif University of Technology	2017-2018 <i>Tehran</i>

TEACHING EXPERIENCE

Teaching Assistant University of California, Irvine	2018-2021 <i>Irvine</i>
Teaching Assistant Sharif University of Technology	2015-2018 <i>Tehran</i>

ABSTRACT OF THE THESIS

Noise Removal using Deep Generative Model

By

Amir Hosein Afandizadeh Zargari

Master of Science in Computer Engineering

University of California, Irvine, 2021

Professor Fadi Kurdahi, Chair

A photoplethysmography (PPG) is a unsophisticated and reasonable-cost optical technique that is frequently utilized in the healthcare field to extract useful health-related data such as heart rate variability, blood pressure, and respiration rate. With the use of portable wearable devices, PPG signals can be captured constantly and remotely. These measuring devices, however, are susceptible to motion artifacts induced by everyday activities. Using extra accelerometer sensors is the most frequent technique to minimize motion artifacts, but they have two drawbacks: (1) Excessive power consumption; (2) the requirement for an accelerometer sensor in a wearable device (which is not required in certain wearables). In this thesis, we provide a non-accelerometer-based PPG motion artifacts reduction method that outperforms previous methods in terms of accuracy. To rebuild clean PPG signals from noisy PPG signals, we employ a Cycle Generative Adversarial Network (CycleGAN). In comparison to the state-of-the-art, our remarkable machine-learning-based technique provides 9.5 times improvement in motion artifact removal without the use of additional sensors like an accelerometer.

Chapter 1

Introduction

A photoplethysmography (PPG) is an uncomplicated, inexpensive, and convenient optical technique used for detecting volumetric blood changes in the microvascular bed of target tissue [2]. Extraction of significant health-related information from PPG signals, ranging from heart rate and heart rate variability to blood pressure and respiration rate, has recently gained more attention in the literature [19].

Nowadays, PPG signals can effortlessly gathered constantly and remotely utilizing low-cost, convenient, and portable wearable devices (e.g., smartwatches, rings, etc.), making them a viable source for everyday wellness applications. PPG signals acquired from portable wearable devices in everyday situations, on the other hand, are frequently captured while a user is engaged in a variety of activities, and hence are distorted by motion distortions. Low signal-to-noise ratio could result in erroneous vital sign extraction, which could have life-threatening repercussions in healthcare applications. Motion artifacts can be detected and removed from PPG signals using a variety of approaches. The majority of efforts on motion artifact detection and filtering in PPG signals fall into one of three categories: (1) non-acceleration based, (2) synthetic reference data, or (3) acceleration data.

The non-acceleration based methods do not require any extra accelerometer sensor for motion artifact detection and removal. Due to the fact that some statistical factors such as skewness and kurtosis would remain unchanged regardless of the presence of noise, these approaches use specific statistical methodologies in previous publications. In [10], such statistical parameters are used to detect and remove the impure parts of the signal caused by motion artifacts. In [7], authors detect motion artifacts using a Variable Frequency Complex Demodulation (VFCDM) method. In this method, the PPG signal is normalized after applying a band-pass filter. Then, to detect motion artifacts, VFCDM distinguishes between the spectral characteristics of noise and clean signals. Then, the unclean-marked signal is deleted from the overall signal due to a frequency shift. In [18], the Discrete Wavelet Transform (DWT) approach is presented as another method in this area.

Because unrecovered noisy data is eliminated from the signal in non-accelerometer based approaches, the clean output signal is usually shorter than the original signal. To mitigate this problem, a synthetic reference signal can be produced from the corrupted PPG signal. In [21], authors use Complex Empirical Mode Decomposition (CEMD) to generate signals. In [11], two PPG sensors are being used to generate a reference signal. One of the sensors is a few millimeters away from the skin, which only measures PPG during movements. First a band-pass filter is applied on both recorded signals; then, an adaptive filter is used to minimize the difference between two recorded signals.

Often an accelerometer sensor is also embedded in wearable devices. To eliminate the effect of motion artifacts, acceleration data can be utilized as a reference signal. In [23], with the help of acceleration data, Singular Value Decomposition (SVD) is used for generating a reference signal for an adaptive filter. After that, an adaptive filter removes motion artifacts from the reference and PPG signals. In a similar approach, the authors of [25] employ a DC remover that employs a different form of adaptive filter. In [6], another method for motion artifact reduction is proposed, which consists of three steps: Signals are windowed,

the resulting signal is filtered, and then a Hankel data matrix is built.

Even though using an accelerometer-based method enhances the model’s accuracy, it suffers from two drawbacks: (1) high power consumption and (2) the requirement to integrate an accelerometer sensor in a wearable device (which is not required in certain wearables). Machine learning techniques can be used as an alternative method to eliminate noise and reconstruct clean signals to address these challenges [5, 26]. Furthermore, machine learning techniques are utilized in healthcare domain in processing of a variety of physiological signals such as PPG for data analysis purposes [3, 4, 15]. The aim of this research is to propose a machine learning non-accelerometer-based PPG motion artifacts removal approach which is low-power and can outperform existing methods in terms of accuracy (even the accelerometer-based techniques).

In recent studies, applying machine learning for image noise reduction has been investigated extensively. The most recent studies use deep generative models to reconstruct or generate clean images [8, 24]. In this study, we propose a novel approach which converts noisy PPG signals to a proper visual representation and uses deep generative models to remove the motion artifacts. To reconstruct clean PPG signals from noisy PPG signals, we employ a Cycle Generative Adversarial Network (CycleGAN) [29]. CycleGAN is a novel and powerful technique in unsupervised learning, which targets learning the distribution of two given datasets to translate an individual input data from the first domain to a desired output data from the second domain. The advantages of CycleGAN over other existing image translation methods are (1) it does not require the pairwise dataset, and (2) the augmentation in CycleGAN makes it practically more suitable for datasets with fewer images. Hence, we use CycleGAN to remove motion artifacts from noisy PPG signals and reconstruct the clean signals. Our experimental results clearly demonstrate the superiority of our approach compared to the current state-of-the-art with a 9.5 times improvement with more energy efficiency due to the elimination of accelerometer sensors.

The remainder of this work is structured as follows. The employed dataset and our proposed pipeline design are described in Section Methods. In the Results section, we summarize the results of our proposed method and compare them to the current state-of-the-art in motion artifact removal from PPG signals. Finally, in the Conclusion section, we address the method's strengths and shortcomings, as well as future research.

It is important to note that this thesis is based on our original paper [27] submitted to Arxiv.org and ACM Health. Thus, it contains information of the our original paper.

Chapter 2

Methods

In this paper, we present an accurate non-accelerometer-based motion artifacts removal model from PPG signals. This model mainly consists of a module for noise detection and another one for motion artifact removal. We present in Figure 2.1 the flow chart of our proposed model. Each module is discussed in detail in the corresponding section.

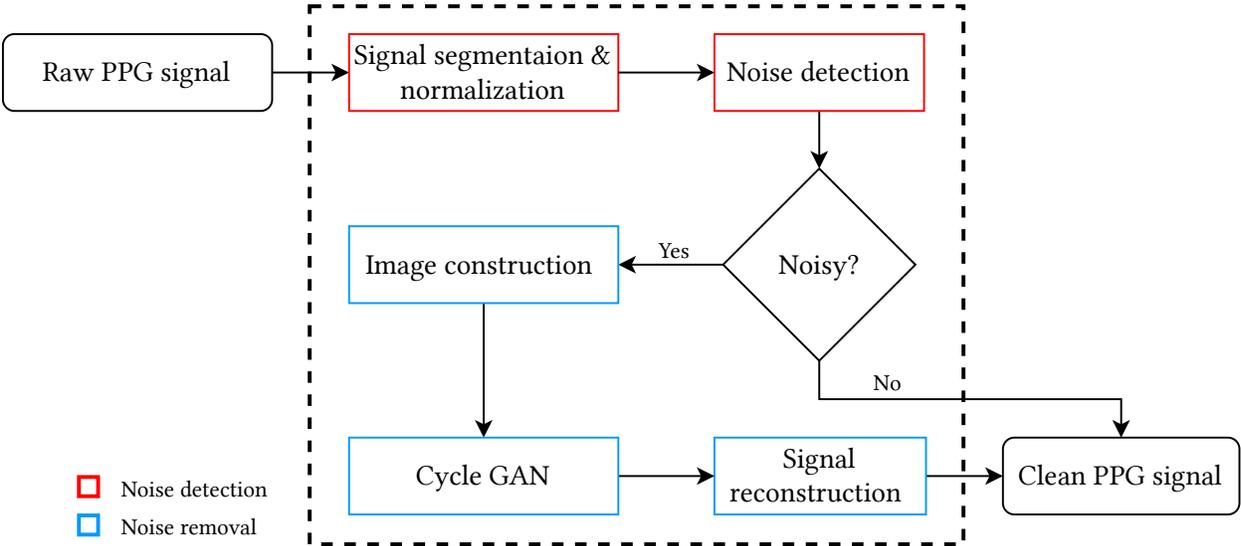


Figure 2.1: Flowchart of the proposed PPG motion artifacts noise removal

In order to train this model, two datasets of PPG signals are required: one consisting of clean

PPG signals and the other one containing noisy PPG signals. The model’s evaluation requires both clean and noisy signals to be taken from the same patient in the same period of time. However, recording such data is not feasible as patients are either performing an activity, which leads to recording a noisy signal or are in a steady-state, which produces a clean signal. For this reason, we simulate the noisy signal by adding data from an accelerometer to the clean signal. This is a common practice and has been used earlier in related work (e.g., [22]) to address this issue. This way, the effectiveness of the model can be evaluated efficiently by comparing the clean signal with the reconstructed output of the model on the derived noisy signal. In the following subsections, we explain the process of data collection for both clean and noisy datasets.

2.1 BIDMC Dataset

For the clean dataset, we use BIDMC dataset [20]. This dataset contains signals and numerics extracted from the much larger MIMIC II matched waveform database, along with manual breath annotations made from two annotators, using the impedance respiratory signal.

The original data was acquired from critically ill patients during hospital care at the Beth Israel Deaconess Medical Centre (Boston, MA, USA). Two annotators manually annotated individual breaths in each recording using the impedance respiratory signal. There are 53 recordings in the dataset, each being 8 minutes long and containing:

- Physiological signals, such as the PPG, impedance respiratory signal, and electrocardiogram (ECG) sampled at 125 Hz.
- Physiological parameters, such as the heart rate (HR), respiratory rate (RR), and blood oxygen saturation level (SpO₂) sampled at 1 Hz.
- Fixed parameters, such as age and gender.

- Manual annotations of breaths.

2.2 Data Collection

We conducted laboratory-based experiments to collect accelerometer data for generating noisy PPG signals. Each of these laboratory-based experiments consisted of 27 minutes of data. A total of 33 subjects participated in the laboratory-based experiments. In each experiment, subjects were asked to perform specific activities while the accelerometer data were collected from them using an Empatica E4 [1] wristband worn on their dominant hand. The Empatica E4 wristband is a medical-grade wearable device that offers real-time physiological data acquisition, enabling researchers to conduct in-depth analysis and visualization. Figure 2.2 shows our experimental procedure. Note that the accelerometer signals are only required for generating/emulating noisy PPG signals, and our proposed motion artifact removal method does not depend on having access to acceleration signals.

Rest	Finger Tapping	Finger Tapping	Rest	Waving	Waving	Rest	Shaking Hands	Shaking Hands
	Low Intensity	High Intensity		Low Intensity	High Intensity		Low Intensity	High Intensity
Activity 1			Activity 2			Activity 3		

Rest	Runnig Arm Swings	Runnig Arm Swings	Rest	Fist Open/Close	Fist Open/Close	Rest	3D Arm Movement	3D Arm Movement
	Low Intensity	High Intensity		Low Intensity	High Intensity		Low Intensity	High Intensity
Activity 4			Activity 5			Activity 6		

Figure 2.2: Experimental procedure to collect accelerometer data.

According to Figure 2.2, each experiment consists of 6 different activities: (1) Finger Tapping, (2) Waving, (3) Shaking Hands, (4) Running Arm Swing, (5) Fist Opening and Closing, and (6) 3D Arm Movement. Each activity lasts 4 minutes in total, including two parts with two different movement intensities (low and high), each of which lasts 2 minutes. Activity

tasks are followed by a 30 seconds rest (R) period between them. During the rest periods, participants were asked to stop the previous activity and put both their arms on a table, and stay in a steady state. Accelerometer data collected during each of the activities were later used to model the motion artifact. We describe this in the next subsection.

2.3 Noisy PPG signal generation

To generate noisy PPG signals from clean PPG signals, we use accelerometer data collected in our study. Clean PPG signals are directly collected from the BIDMC dataset. We down-sample these signals to 32 Hz to ensure they are synchronized with the collected accelerometer data.

Empatica has an onboard MEMS type 3-axis accelerometer that measures the continuous gravitational force (g) applied to each of the three spatial dimensions (x, y, and z). The scale is limited to $\pm 2g$. Figure 2.3 shows an example of accelerometer data collected from E4.

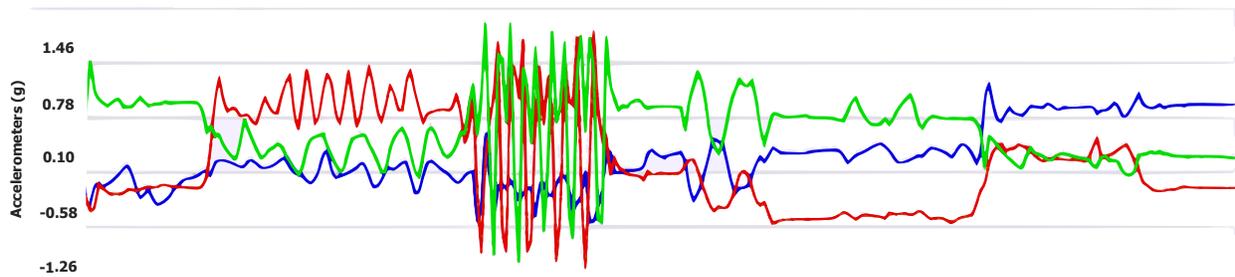


Figure 2.3: An example of Accelerometer data from Connect, the subject moves into position, walks, runs, and then simulates the turning of a car’s steering wheel. The dimensional axes are depicted in red, green and blue.

Along with the raw 3-dimensional acceleration data, Empatica also provides a moving average of the data. Figure 2.4 visualizes the moving averaged data.



Figure 2.4: The same data as Figure 2.3 is visualized using the moving average. From Connect, the subject moves into position, walks, runs, and then simulates the turning of a car’s steering wheel. The dimensional axes are depicted in red, green and blue.

The following formula is used to calculate the moving averaged of the data,

$$\text{Sum+} = \max(|\text{Curr}_X - \text{Prev}_X|, |\text{Curr}_Y - \text{Prev}_Y|, |\text{Curr}_Z - \text{Prev}_Z|) \quad (2.1)$$

in which, Curr_i and Prev_i are respectively the current value and the previous value of the accelerometer sensor (g) along the i -th dimension. The $\max(a, b, c)$ function returns the maximum value among a , b , and c .

Then the following formula is used to filter the output:

$$\text{Avg} = 0.9 \times \text{Avg} + 0.1 \times \frac{\text{Sum}}{32} \quad (2.2)$$

The filtered output (Avg) is directly used as a model for motion artifacts in our study. To simulate the noisy PPG signals, we add this noise model to a 2 minutes window of the clean PPG signals collected from the BIDMC dataset. We use 40 out of 53 signals in BIDMC directly as the clean dataset for training. Among these 40 signals, 20 are selected and augmented with the accelerometer data to construct the noisy dataset for training. The 13 remaining BIDMC signals and accelerometer data were added together to form the clean and noisy datasets for testing. In the rest of this section we describe each part of the model

introduced in Figure 2.1.

2.4 Noise Detection

To perform noise detection, first, the raw signal is normalized by a linear transformation to map its values to the range (0, 1). This can be performed using a simple function as below:

$$\text{Sig}_{\text{norm}} = \frac{\text{Sig}_{\text{raw}} - \min(\text{Sig}_{\text{raw}})}{\max(\text{Sig}_{\text{raw}}) - \min(\text{Sig}_{\text{raw}})} \quad (2.3)$$

where Sig_{raw} is the raw signal and Sig_{norm} is the normalized output. Then, the normalized signal is divided into equal windows of size 256, which is the same window size we use for noise removal. These windows are then used as the input of the noise detection module to identify the noisy ones.

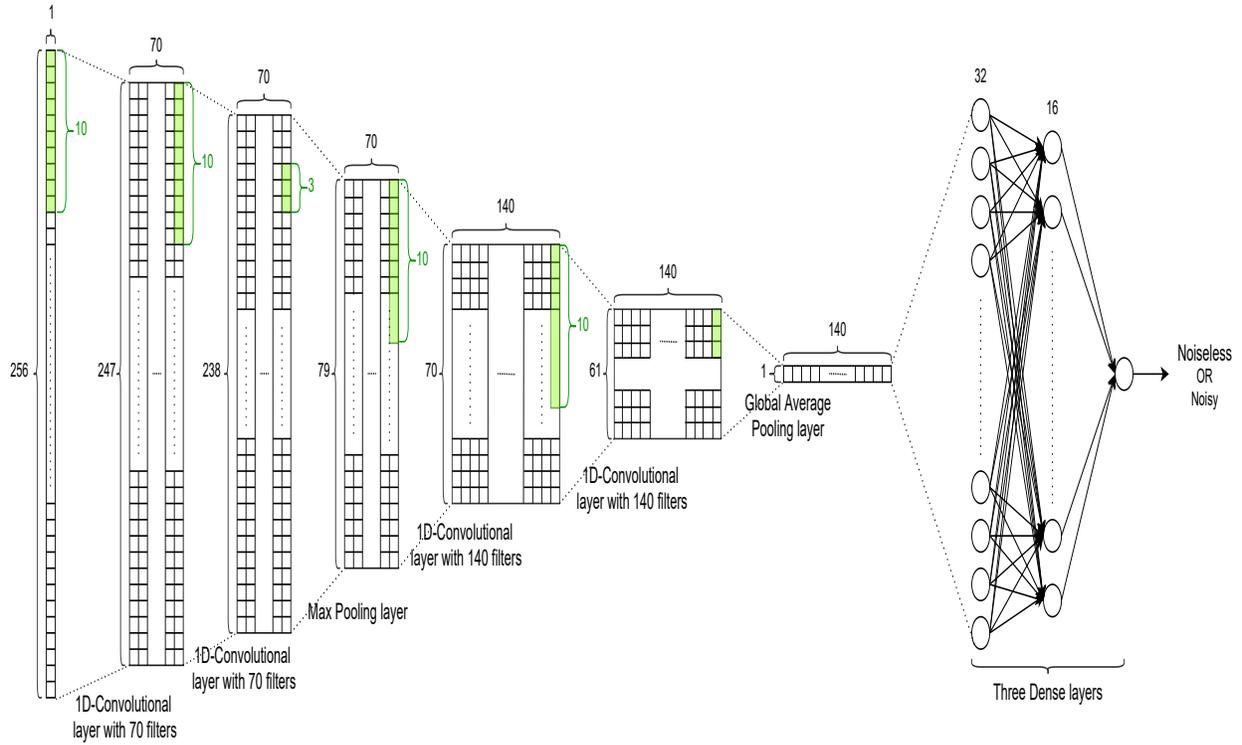


Figure 2.5: The structure of the noise detection model.

The similar type of machine learning network used in [28] can be employed as a noise detection system. To explain the network structure for the noise detection method (Table 2.1 and Figure 2.5), first, we use a 1D-convolutional layer with 70 initial random filters with a size of 10 to select the basic features of the input data and convert the matrix size from 256×1 to 247×70 . To extract more complex features from the data, another 1D-convolutional layer with the same filter size 10 is required. As the third layer, a pooling layer with a filter size of 3 is utilized. In this layer, a sliding window slides over the input of the layer and in each step, the maximum value of the window is applied to the other values. This layer converts a matrix size of 238×70 to 79×70 . To select additional complex features, another set of convolutional layers are used with a different filter size. This set is followed by two fully connected layers of sizes 32 and 16. Lastly, a dense layer of size 2 with a softmax activation would produce the probability of each class: clean and noisy. The maximum of these two probabilities would be identified as the result of the classification. The accuracy of our proposed binary classification method is 99%, which means that the system can almost always detect a noisy signal from a clean signal.

Table 2.1: The layer configuration of the noise detection model.

Layer	Structure	Output
Conv1D+Relu	70×10	247×70
Conv1D+Relu	70×10	238×70
Max pooling 1D	3	79×70
Conv1D+Relu	140×10	70×140
Conv1D+Relu	140×10	61×140
Global average pooling	N/A	140
Dense+Relu	128	32
Dense+Relu	16	16
Dense+Softmax	2	2

2.5 Noise Removal

In this section, we explore the reconstruction of noisy PPG signals using deep generative models. Once a noisy window is detected, it is sent to the noise removal module for further processing. First, the windows are transformed into 2-dimensional images, to exploit the power of existing image noise removal models, and then a trained CycleGAN model is used to remove the noise induced by the motion artifact from these images. In the final step of the noise removal, the image transformation is reversed to obtain the clean output.

The transformation needs to provide visual features for unexpected changes in the signal so that the CycleGAN model would be able to distinguish and hence reconstruct the noisy parts. To extend the 1-dimensional noise on the signal into a 2-dimensional visual noise on the image, we apply the following transformation:

$$\text{Img}_{i,j} = \text{floor}((\text{Sig}[i] + \text{Sig}[j]) \times 128) \quad (2.4)$$

where Sig is a normalized window of the signal. Each pixel will then have a value between 0 and 255, representing a grayscale image.

An example of such transformation is provided in Figure 2.6 for both the clean and the noisy signal. According to this figure, the noise effect is visually observable in these images.

Autoencoders and CycleGAN are two of the most powerful approaches for image translation. These methods have proven to be effective in the particular case of noise reduction. Autoencoders require the pairwise translation of every image in the dataset. In our case, clean and noisy signals are not captured simultaneously, and their quantity differs. CycleGAN, on the other hand, does not require the dataset to be pairwise. Also, the augmentation in CycleGAN makes it practically more suitable for datasets with fewer images. Hence, we use

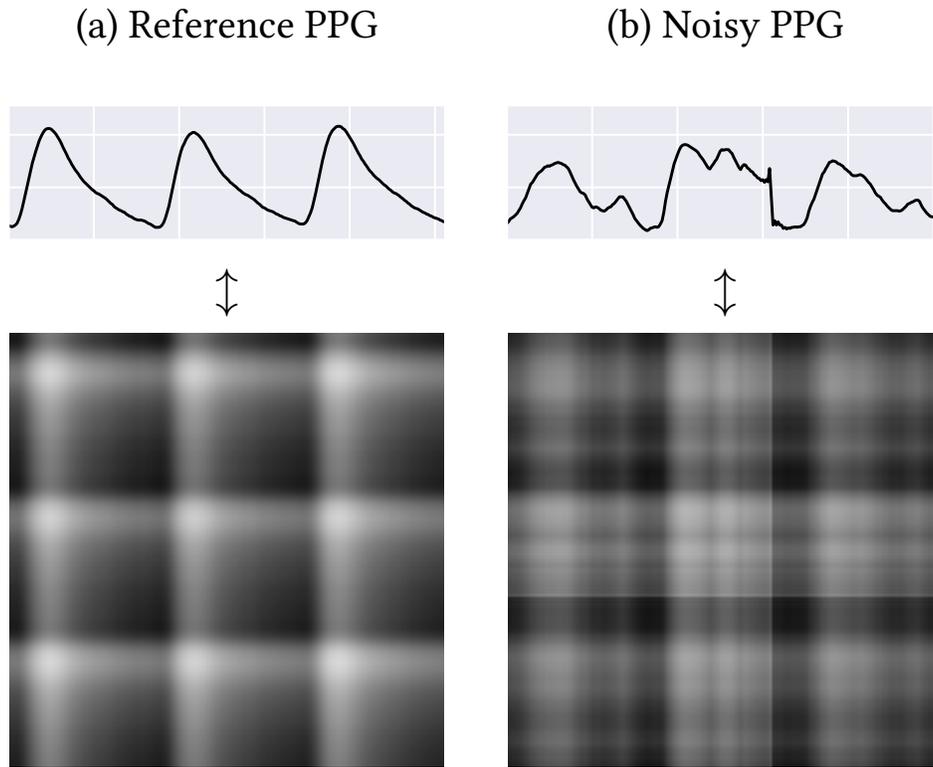


Figure 2.6: An example of signal to image transformation.

CycleGAN to remove motion artifacts from noisy PPG signals and reconstruct the clean signals.

CycleGAN is a Generative Adversarial Network designed for the general purpose of image-to-image translation. CycleGAN architecture was first proposed by Zhu et al. in [29].

The GAN architecture consists of two networks: a generator network and a discriminator network. The generator network starts from a latent space as input and attempts to generate new data from the domain. The discriminator network aims to take the generated data as an input and predict whether it is from a dataset (real) or generated (fake). The generator is updated to generate more realistic data to better fool the discriminator, and the discriminator is updated to better detect generated data by the generator network.

The CycleGAN is an extension of the GAN architecture. In the CycleGAN, two generator

networks and two discriminator networks are simultaneously trained. The generator network takes data from the first domain as an input and generates data for the second domain as an output. The other generator takes data from the second domain and generates the first domain data. The two discriminator networks are trained to determine how plausible the generated data are. Then the generator models are updated accordingly. This extension itself cannot guarantee that the learned function can translate an individual input into a desirable output. Therefore, the CycleGAN uses a cycle consistency as an additional extension to the model. The idea is that output data by the first generator can be used as input data to the second generator. Cycle consistency is encouraged in the CycleGAN by adding an additional loss to measure the difference between the generated output of the second generator and the original data (and vice versa). This guides the data generation process toward data translation.

In our CycleGAN architecture, we apply adversarial losses [9] to both mapping functions ($G : X \rightarrow Y$ and $F : Y \rightarrow X$). The objective of the mapping function G as a generator and its discriminator D_Y is expressed as below:

$$L_{GAN}(G, D_Y, X, Y) = E_{y \sim p_{data}(y)}[\log \log D_Y(y)] + E_{x \sim p_{data}(x)}[\log \log(1 - D_Y(G(x)))] \quad (2.5)$$

where the function G takes an input from domain X (e.g., noisy PPG signals), attempting to generate new data that look similar to data from domain Y (e.g., clean PPG signals). In the meantime, D_Y aims to determine whether its input is from the translated samples $G(x)$ (e.g., reconstructed PPG signals) or the real samples from domain Y . A similar adversarial loss is defined for the mapping function $F : Y \rightarrow X$ as $L_{GAN}(F, D_X, Y, X)$.

As discussed before, adversarial losses alone cannot guarantee that the learned function can map an individual input from domain X to the desired output from domain Y . In [29], the authors argue that to reduce the space of possible mapping functions even further, learned

mapping functions (Y and F) need to be cycle-consistent. This means that the translation cycle needs to be able to translate back the input from domain X to the original image as $X \rightarrow G(X) \rightarrow F(G(X)) \sim X$. This is called forward cycle consistency. Similarly, backward cycle consistency is defined as: $y \rightarrow F(y) \rightarrow G(F(y)) \sim y$. This behavior is presented in our objective function as:

$$L_{\text{cyc}}(G, F) = E_{x \sim p_{\text{data}}(x)}[\|F(G(x)) - x\|_1] + E_{y \sim p_{\text{data}}(y)}[\|G(F(y)) - y\|_1] \quad (2.6)$$

Therefore, the final objective of CycleGAN architecture is defined as:

$$L(G, F, D_X, D_Y) = L_{\text{GAN}}(G, D_Y, X, Y) + L_{\text{GAN}}(F, D_X, Y, X) + \lambda L_{\text{cyc}}(G, F) \quad (2.7)$$

where λ controls the relative importance of the two objectives.

In Equation 2.7, G aims to minimize the objective while an adversary D attempts to maximize it. Therefore, our model aims to solve:

$$G^*, F^* = \operatorname{argmin} L(G, F, D_X, D_Y) \quad (2.8)$$

The architecture of the generative networks is adopted from Johnson et al. [14]. This network contains four convolutions, several residual blocks [12], and two fractionally-strided convolutions with stride 0.5. For the discriminator networks, they use 70×70 PathGANs [17, 13, 16].

After the CycleGAN is applied to the transformed image, the diagonal entries are used to retrieve the reconstructed signal.

$$\text{Sig}_{\text{rec}}[i] = \text{Img}[i, i]/256 \quad (2.9)$$

Chapter 3

Results

In this section, we assess the efficiency of our model based on the following measures: root mean square error (RMSE) and peak-to-peak error (PPE). A signal window size of 256 and an image size of 256 by 256 were used for all experimental purposes, and 25% of the data was assigned for validation. The noise detection module had an accuracy of 99%. The summary of the results for noise removal, including the improvement for each noise type and noise intensity, can be found in Table 3.1.

For each noise type, there are two entries in this table, one corresponding to the slow movement and the other one corresponding to the fast movement. The average S/N value for slow movements is 21.7dB, as provided in the table, while the average S/N value for fast movements is 14.0dB. For each of the measures, RMSE and PPE, we calculated the error between the generated signal and the reference signal as well as the error between the noisy signal and the reference signal in order to observe the improvement of the model on the noisy signal. The degree of improvement on each noise type is added in a separate column in the table. According to the table, the average of improvement on RMSE is $41\times$ and the average of improvement on PPE is $58\times$.

Table 3.1: Results of the proposed method.

Noise Type	S/N (dB)	RMSE Gen. (BPM)	RMSE Nsy. (BPM)	RMSE Imprv.	PPE Gen. (BPM)	PPE Nsy. (BPM)	PPE Imprv.
Waving	20.04	0.213	41.76	196.07×	0.136	32.89	241.60×
Waving	11.30	2.43	55.30	22.75×	1.088	37.90	34.84×
3D Arm Movement	20.17	1.644	92.12	56.03×	0.772	44.03	57.06×
3D Arm Movement	13.12	1.688	65.99	39.10×	0.700	48.49	69.29×
Shaking Hands	21.66	1.556	61.89	39.78×	0.576	28.62	49.71×
Shaking Hands	14.96	4.203	84.31	20.06×	2.677	64.58	24.12×
Finger Tapping	22.99	1.758	63.43	36.07×	0.653	45.14	69.14×
Finger Tapping	13.99	3.008	21.76	7.235×	1.191	10.70	8.99×
Fist Open Close	25.11	1.648	35.74	21.69×	0.528	24.51	46.44×
Fist Open Close	16.69	2.151	51.28	23.84×	1.113	42.65	38.33×
Running Arm	20.14	2.056	22.93	11.16×	0.715	19.32	27.02×
Running Arm	13.98	3.807	77.73	20.42×	1.348	50.75	37.64×
Average	17.85	2.18	56.19	41.18×	0.958	37.465	58.68×

An example of a reconstructed signal is presented in Figure 3.1, together with the noisy PPG and the reference PPG signal. As we can see in this figure, the noise is significantly reduced, and the peak values are adjusted accordingly, confirming that the image transformation successfully represents the noise in a visual format.

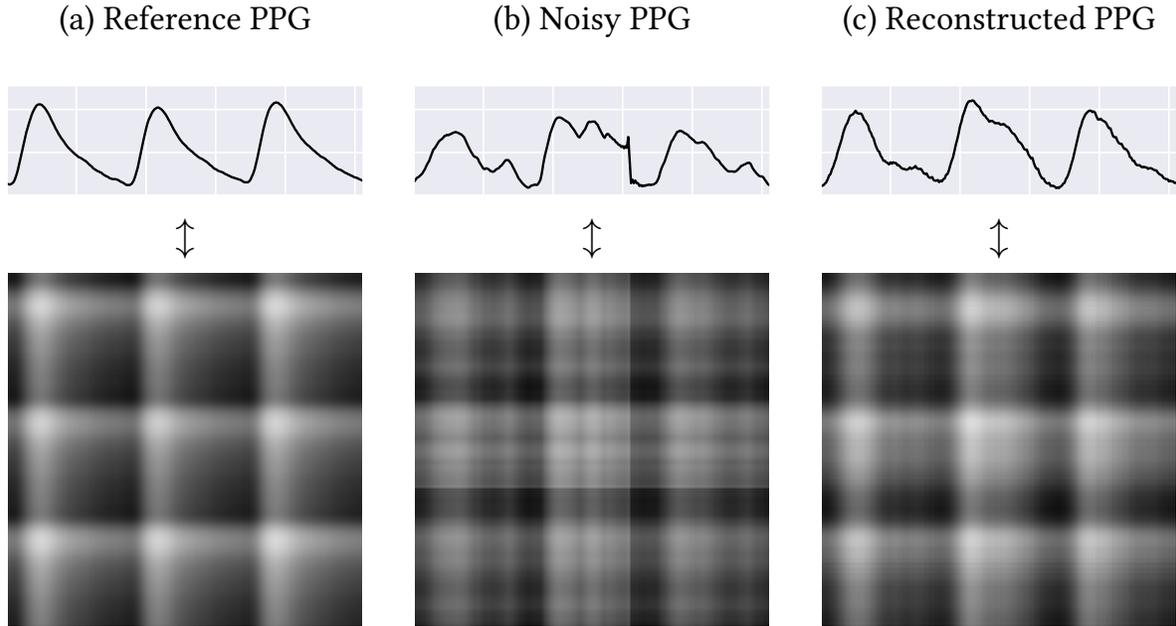


Figure 3.1: The reconstructed signal of Figure 2.6 alongside with the noisy and the reference signals

3.1 Comparison

In this section we compare our model’s efficiency with the state-of-the-art (Table 3.2). To minimize the difference between our experimental setup and the setups used in the related works we use the same measures. It should be noted that it is not feasible to perform a close comparison between our model and the existing works, due to the differences in the dataset and the lack of a public dataset providing noisy and clean signals simultaneously.

Table 3.2: The summary comparison of our result with the existing methods. MAE stands for Mean absolute error.

Paper	Method	Accelerometer	Before	Outcome
Proposed method	CycleGAN	No	PPE 37.46 BPM RMSE 56.18 BPM	PPE 0.95 BPM RMSE 2.18 BPM
Hanyu and Xiaohui [10]	Statistical Evaluation	No	PPE 8.1 BPM	PPE 7.85 BPM
Bashar et al. [7]	VFCDM	No	N/A	6.45% false positive
Lin and Ma [18]	DWT	No	PPE 13.97 BPM	PPE 6.87 BPM
Raghuram et al. [21]	CEMD LMS	Syn.	PPE 0.466 BPM	PPE 0.392 BPM
Hara et al. [11]	NLMS and RLS	Syn.	RMSE 28.26 BPM	RMSE 6.5 BPM
Tanweer et al. [23]	SVD and X-LMS	Yes	N/A	PPE 1.37 BPM
Wu et al. [25]	DC remover and RLS	Yes	N/A	STD 3.81
Bac’a et al. [6]	MAR and AT	Yes	N/A	MAE 2.26 BPM

In comparison to non-accelerometer-based methods, our model significantly outperforms these models. The best performance observed in previous work is reported in [11] that improves the average RMSE from 28.26BPM to 6.5BPM ($4.3\times$ improvement). However, our model’s improvement on average RMSE is from 56.18 to 2.18 ($25.8\times$ improvement). In most of the existing accelerometer-based methods, no value is provided for the degree of the input noise. Although the best reported PPE belongs to [21] with an outcome of 0.392BPM, the best improvement is achieved by [18] from 13.97BPM to 6.87BPM ($2.03\times$ improvement). However, our model’s improvement on average PPE is from 37.46BPM to 0.95BPM ($39.4\times$ improvement).

Chapter 4

Discussion

Noise reduction has been extensively studied in image processing, and the introduction of powerful models such as CycleGAN has shown promising results in terms of noise reduction in images. Inspired by this fact, we proposed a signal to image transformation that visualizes signal noises in the form of image noise. To the best of our knowledge, this is the first use of CycleGAN for bio-signal noise reduction which eliminates the need for an accelerometer to be embedded into wearable devices, which in turn helps to reduce the power consumption and cost of these devices.

It should be noted that despite the significant benefits of our proposed method in removing noise in different situations, it may not be effective in all possible scenarios. Clearly, the intensity of noise applied to the signals, and the variations of the noise, also called noise categories, are controlled for the purpose of this study. When the signal is faded in the noise, this method may not be applicable. Although it will improve the error, it does not guarantee a reasonable upper bound. However, the same limitations also exist in the related works.

Chapter 5

Conclusions

In this study, we introduced an image processing approach to the problem of noise removal from PPG signals where the noise is selected from a set of noise categories that simulate the daily routine of a person. This method does not require an accelerometer on the sensor, therefore, it can be applied to other variations of physiological signals, such as ECG, to reduce the power usage of the measuring device and improve its efficiency. In this work, the novel use of CycleGAN as an image transformer is leveraged to transform such physiological signals. On average, the reconstructed PPG performed using our proposed method offers $41\times$ improvement on RMSE and $58\times$ improvement on PPE, outperforming the state-of-the-art by a factor of 9.5.

Bibliography

- [1] Empatica — medical devices, ai and algorithms for remote patient monitoring. <https://www.empatica.com/>. Accessed: 2021-05-24.
- [2] J. Allen. Photoplethysmography and its application in clinical physiological measurement. *Physiological measurement*, 28(3):R1, 2007.
- [3] S. A. H. Aqajari, R. Cao, E. K. Naeini, M.-D. Calderon, K. Zheng, N. Dutt, P. Liljeberg, S. Salanterä, A. M. Nelson, and A. M. Rahmani. Pain assessment tool with electrodermal activity for postoperative patients: Method validation study. *JMIR mHealth and uHealth*, 9(5):e25258, 2021.
- [4] S. A. H. Aqajari, E. K. Naeini, M. A. Mehrabadi, S. Labbaf, N. Dutt, and A. M. Rahmani. pyeda: An open-source python toolkit for pre-processing and feature extraction of electrodermal activity. *Procedia Computer Science*, 184:99–106, 2021.
- [5] M. Ashrafiamiri, S. Manoj Pudukotai Dinakarrao, A. H. Afandizadeh Zargari, M. Seo, F. Kurdahi, and H. Homayoun. R2ad: Randomization and reconstructor-based adversarial defense on deep neural network. In *Proceedings of the 2020 ACM/IEEE Workshop on Machine Learning for CAD*, pages 21–26, 2020.
- [6] A. Baca, G. Biagetti, M. Camilletti, P. Crippa, L. Falaschetti, S. Orcioni, L. Rossini, D. Tonelli, and C. Turchetti. Carma: a robust motion artifact reduction algorithm for heart rate monitoring from ppg signals. In *2015 23rd European signal processing conference (EUSIPCO)*, pages 2646–2650. IEEE, 2015.
- [7] S. K. Bashar, D. Han, A. Soni, D. D. McManus, and K. H. Chon. Developing a novel noise artifact detection algorithm for smartphone ppg signals: Preliminary results. In *2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, pages 79–82. IEEE, 2018.
- [8] J. Chen, J. Chen, H. Chao, and M. Yang. Image blind denoising with generative adversarial network based noise modeling. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3155–3164, 2018.
- [9] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial networks. *arXiv preprint arXiv:1406.2661*, 2014.

- [10] S. Hanyu and C. Xiaohui. Motion artifact detection and reduction in ppg signals based on statistics analysis. In *2017 29th Chinese control and decision conference (CCDC)*, pages 3114–3119. IEEE, 2017.
- [11] S. Hara, T. Shimazaki, H. Okuhata, H. Nakamura, T. Kawabata, K. Cai, and T. Takubo. Parameter optimization of motion artifact canceling ppg-based heart rate sensor by means of cross validation. In *2017 11th international symposium on medical information and communication technology (ISMICT)*, pages 73–76. IEEE, 2017.
- [12] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.
- [13] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1125–1134, 2017.
- [14] J. Johnson, A. Alahi, and L. Fei-Fei. Perceptual losses for real-time style transfer and super-resolution. In *European conference on computer vision*, pages 694–711. Springer, 2016.
- [15] K. Joshi, A. Javani, J. Park, V. Velasco, B. Xu, O. Razorenova, and R. Esfandyarpour. A machine learning-assisted nanoparticle-printed biochip for real-time single cancer cell analysis. *Advanced Biosystems*, 4(11):2000160, 2020.
- [16] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, et al. Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4681–4690, 2017.
- [17] C. Li and M. Wand. Precomputed real-time texture synthesis with markovian generative adversarial networks. In *European conference on computer vision*, pages 702–716. Springer, 2016.
- [18] W.-J. Lin and H.-P. Ma. A physiological information extraction method based on wearable ppg sensors with motion artifact removal. In *2016 IEEE international conference on communications (ICC)*, pages 1–6. IEEE, 2016.
- [19] M. A. Mehrabadi, S. A. H. Aqajari, I. Azimi, C. A. Downs, N. Dutt, and A. M. Rahmani. Detection of covid-19 using heart rate and blood pressure: Lessons learned from patients with ards, 2020.
- [20] M. A. Pimentel, A. E. Johnson, P. H. Charlton, D. Birrenkott, P. J. Watkinson, L. Tarassenko, and D. A. Clifton. Toward a robust estimation of respiratory rate from pulse oximeters. *IEEE Transactions on Biomedical Engineering*, 64(8):1914–1923, 2016.

- [21] M. Raghuram, K. Sivani, and K. A. Reddy. Use of complex emd generated noise reference for adaptive reduction of motion artifacts from ppg signals. In *2016 international conference on electrical, electronics, and optimization techniques (ICEEOT)*, pages 1816–1820. IEEE, 2016.
- [22] M. S. Roy, R. Gupta, J. K. Chandra, K. D. Sharma, and A. Talukdar. Improving photoplethysmographic measurements under motion artifacts using artificial neural network for personal healthcare. *IEEE Transactions on Instrumentation and Measurement*, 67(12):2820–2829, 2018.
- [23] K. T. Tanweer, S. R. Hasan, and A. M. Kamboh. Motion artifact reduction from ppg signals during intense exercise using filtered x-lms. In *2017 IEEE international symposium on circuits and systems (ISCAS)*, pages 1–4. IEEE, 2017.
- [24] L. D. Tran, S. M. Nguyen, and M. Arai. Gan-based noise model for denoising real images. In *Proceedings of the Asian Conference on Computer Vision*, 2020.
- [25] C.-C. Wu, I.-W. Chen, and W.-C. Fang. An implementation of motion artifacts elimination for ppg signal processing based on recursive least squares adaptive filter. In *2017 IEEE biomedical circuits and systems conference (BioCAS)*, pages 1–4. IEEE, 2017.
- [26] R. Yasaei, F. Hernandez, and M. A. Al Faruque. Iot-cad: context-aware adaptive anomaly detection in iot systems through sensor association. In *2020 IEEE/ACM International Conference On Computer Aided Design (ICCAD)*, pages 1–9. IEEE, 2020.
- [27] A. H. A. Zargari, S. A. H. Aqajari, H. Khodabandeh, A. M. Rahmani, and F. Kurdahi. An accurate non-accelerometer-based ppg motion artifact removal technique using cyclegan. *arXiv preprint arXiv:2106.11512*, 2021.
- [28] A. H. A. Zargari, M. Dautta, M. Ashrafiamiri, M. Seo, P. Tseng, and F. Kurdahi. Newertrack: Ml-based accurate tracking of in-mouth nutrient sensors position using spectrum-wide information. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 39(11):3833–3841, 2020.
- [29] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, pages 2223–2232, 2017.