Respiratory Rate Estimation on Measurements of Wearable Device

Federico Zanotti†

***Abstract*—Breathing rate (BR) is a key physiological parameter used in a range of clinical settings. Despite its diagnostic and prognostic value, it is still widely measured by counting breaths manually. The modulations presented in the photoplethysmogram (PPG) have been useful to derive RR using signal processing. A Deep Learning framework, based on the Convolutional Neural Networks (CNN), is proposed to estimate RR using PPG. This approach takes into account also data recorded from the Oxyrecord, a professional pulse oximiter. A synthetic dataset is used to increment the size of data available. In the end, the best result is obtained with a Mean Absolute Error (MAE) of 2.84 ± 0.54 using 5-Fold Cross Valdation. The model achieved good perfomance and also it was evaluated on PPG signal measured by the wearable device. This encourages further research on estimating RR through the signal measured by a wearable device.**

***Index Terms*—Respiratory Rate, PhotoPlethysmoGram, Convolutional Neural Network, Pulse Oximiter**

I. Introduction

Vital signs are used to measure body's basic functions. These measurements can give some important information of the physical health of a person or can discover diseases that can be predicted before they happen. This last point is very important, because if a disease is discovered in his initial forms, then it can be faced immediately reducing the possibility to reach his evolution to more difficult forms. One important Vital Sign is the Respiratory Rate (RR): the number of breaths a person takes in one minute. An abnormal Respiratory Rate can predict cardiac arrest and it's highly correlated with in-hospital mortality [1]. Moreover, several studies showed that assess the respiratory rate in high-risk COVID-19 patients could accelerate life-saving treatments [2]. Consequently, monitoring RR is necessary to assess patient's health both in hospital and at home. Nowadays the clinical practice for the RR estimation is done by counting chest wall expansions, but it's very inconvient to assess if a person has an abnoral RR and mostly it's not continuos. This necessitates the use of unobtrusive sensors for obtaining respiratory information from patients in normal situations. In the last years the research on wearable medical devices has taken huge forward steps, allowing to continuosly monitoring key phisiological measurements like heart rate, and temperature rate. Many smartwatches and wearable devices allow for an ambulatory monitoring of the photoplethysmogram (PPG) signal. The PPG is modulated predominantly by respiratory system [3] and so it can be used to predict Respiratory Rate. In particular PPG presents three types of modulation: baseline wander (BW) of the PPG that is influenced by the changes in intrathoracic pressure and vasoconstriction of arteries during inhalation; amplitude modulation (AM) of the PPG that reflects the changes in stroke volume and intrathoracic pressure during respiratory cycles; and frequency modulation (FM), the respiratory sinus arrhythmia that exhibits HR to increase during inspiration and to decrease during expiration. A lot of research has done to estimate RR based on PPG processing, with time and or frequency domain analysis, signal decomposition and modulations. More recent advantages are obtained using Deep Learning models, expecially with the Convolutional Neural Networks (CNN) commonly used for images but can be applied also to time series like a signal. In particular this project enriches the work done in [4], with the contribution of data recorded by a wearable device called the Oxyrecord, a professional pulse oximiter. Moreover in this project more types of synthetic data are used showing how a good synthetic signal can be generated. In summary the contributions in the paper are as follows:

- I propose a full CNN on PPG signal to predict respiratory rate
- I tested the network with combinations of synthetic datasets and real dataset available online
- I fine tuned the network on recorded PPG signal via wearable sensor

II. Related Work

In the literature there are many algorithms that estimate RR using the PPG signal and a nice summary is descripted in [5] where the algorithms are based on digital filtering, signal decomposition or they focus on the modulations of the PPG signal. These classical approaches are based on handcrafted rules and empirical parameters, while Deep Learning has recently been used for biometric analysis. In [6] different Deep Learning algorithms are used to classify hypertensive into no hypertension, prehypertension, stage I hypertension, and stage II hypertension while in [7] Deep Learning is used for predicting blood pressure from PPG signals. There are different works in which the respiration rate is estimated using a Deep Learning approach. In [8] Cycle Generative Adversarial Networks are used while in [9] different Machine Learning algorithms are employed. All the works citated has used for training and testing the algorithms data that comes from publicy available dataset, while in this work the model

†Department of Mathematics, University of Padova, email: federico.zanotti@studenti.unipd.it

Special thanks to Francesca Meneghello and Silvia Zampato

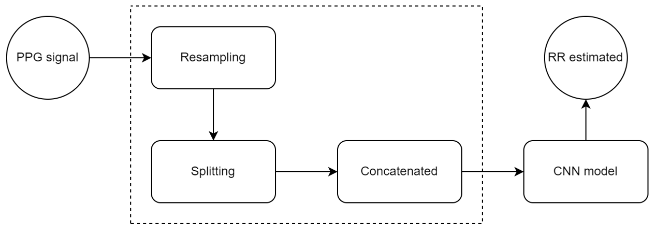

Fig. 1: Processing Pipeline

used is tested on real data collected by a wearable sensor. A mentions must be done on the paper [10], in which they developed an encoder decoder structure using dilated residual block. Their results are very significant but they are tested on two common public available dataset.

III. Processing Pipeline

In this section a synthetic processing pipeline of the project is described. First of all the PPG signals are resampled to the frequency of the pulse oximiter. Secondly, all the signals are splitted in single signals of 60 seconds long. This choice is supported by [11], in which the optimal windows size for RR estimation can range from 30 seconds to 90 seconds with lower error with larger windows size. In order to reduce the computational cost, a window size of 60 seconds is used. Then all signals are standardized with the formula 1 with respect to the signals recorded by the Oxyrecord.

$$z = \frac{x - \mu}{s}. \tag{1}$$

The signals that act as input are then concatenated together and given as raw data to the CNN model. Finally, the model returns the predicted respiration rate as a decimal number. This entire process is visualized in Fig. 1

IV. Signals and Features

The signals employed are real data that comes from available medical dataset, synthetic PPG values generated and data recorded by the pulse oximiter.

Synthetic Signals: Synthetic PPG signals are generated under the influence of of three respiratory modulations: baseline wander (BW), amplitude modulation (AM), and frequency modulation (FM). Each recording is reproduced for 300 seconds with the sample rate equal to the frequency of the sensor (96.68). The RR input for synthetic data generation was chosen from 4 to 60 breaths-per-min (brpm).

In this work different type of synthetic data are generated by varying the three respiratory modulations. The code used for the generation is gently provided by [12]

Real signal: the real signals come from BIDMC, a dataset extracted from the MIMIC-II resource [13]. The original data was acquired from critically-ill patients during hospital care at the Beth Israel Deaconess Medical Centre (Boston, MA, USA). BIDMC is composed by PPG recordings from 53 critically-ill patients of 8-minute duration with sampling rate of $125Hz$. Two annotators manually annotated individual breaths in each recording using the impedance respiratory signal

Data Recorded: The data recorded is obtained by the Oxyrecord. This device is a professional pulse oximiter for real time and long recordings. This wearable device measures the oxigen saturation (SpO2), that corresponds to the amount of oxygen-carrying hemoglobin in the blood relative to the amount of hemoglobin not carrying oxygen. Moreover it computes the pulse rate and through the mobile application Remocop, provided by the University of Padua, PPG values for every timestamps of the recordings are retrieved. The measurements are done on a healthy male person of 25 years old. The number of breath are taken in a manual way counting the number of times the chest of the subject rises. In particular to avoid a situation in which the person is influenced by the knowledge that his breath is measured, a distracting activity is presented to the patient. The process is recorded also via video, in order to count the number of breaths through the chest movement.

All these signals are represented with a vector of three dimensions. The first dimension defines the number of samples, the second defines the number of features and the third is added in order to be consistent with the input required by the CNN. Moreover the synthetic and BIDMC data are then splitted in training and validation set, while the model is evaluated on the data recorded.

V. Learning Framework

A. Dataset

The dataset used for training is a mixture of the BIDMC and synthetic data. in particular synthetic data are modulated by changing the amplitude of the AM, BW, FM and a generic modulation of amplitude of all data (General Amplitude). In the Table 1 is listed all the different combinations obtained.

B. Model

In this project a CNN architecture is used to determine the RR given in input the PPG signal. In particular the CNN uses the residual blocks, a common variant of the CNN: ResNet [14]. This one inserts skip connections in order to jump over some layers and avoiding the problem of vanishing of gradient. In the current experiment, the residual blocks are employed with five convolutional layers as depicted in Fig 2.

TABLE 1: Different Synthetic Data

Modulations	**Magnitude**
AM and BW	0.1
AM and BW	0.2
AM and BW	0.3
AM and BW	0.4
FM	0.05
FM	0.03
FM	0.025
General amplitude	1
General amplitude	3
General amplitude	6

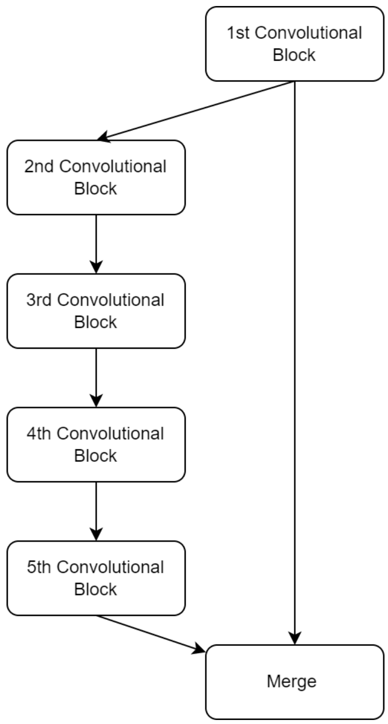

Fig. 2: Residual Block

The model is built with five residual blocks, followed by an Average Pooling layer that reduces the dimensionality. Then a Flatten layer is used in order to have a vector-like representation and finally two fully connected layer plus the output layer conclude the network (Figure 3). The activation function used for all the layer is the ReLU, a common non linear function used in Machine Learning that allows to learn non-linear input. Finally the network uses Adam as

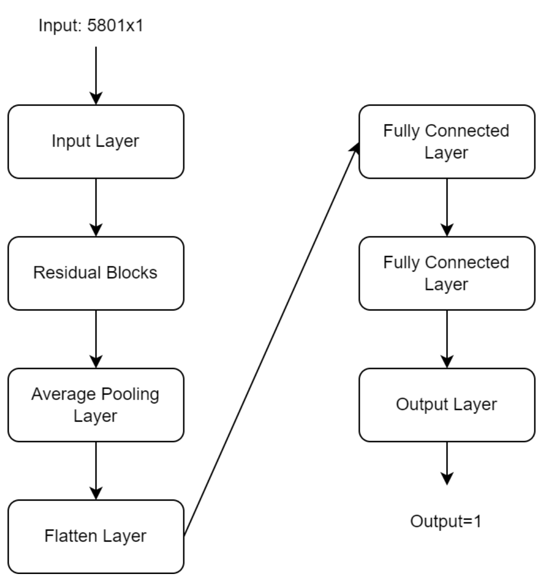

Fig. 3: Residual Block

optimization algorithm. As a loss Mean Absolute Error (MAE) is used in order also to compare the results. MAE (formula 2) is a performance metric that averages the absolute difference between the real RR and the one estimated by the model.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|RR^{i}_{estimated} - RR^{i}_{true}| \quad (2)$$

The training is done using 5-fold cross validation, a resampling procedure used to evaluate machine learning models on a limited data sample. The training is done for 50 epochs with a batch size of 254. During the training a tecnique called Early Stopping is used to prevent overfitting, and a patience parameter of 5 is used on the validation loss. This means that if the validation loss does not improve for the last 5 epochs, than the training stops and the best model with the lowest validations loss is returned

VI. Results

A crucial step in Deep Learning models is the tuning of hyperparameter. In this work this is done via a manual search and the model is fine tuned on the recorded data, in order to maximize the precision of the estimation given by the pulse oximiter. The hyperparameters for the first residual block are descripted in in Table 2 while in the remaining residual blocks the number of filters is set to 6 and the other parameters are keep fixed. For the the rest of the network in the Table 3 there are the final parameters values.

TABLE 2: Residual Block Hyperparams

Layer	**Value**
1^{st} Convolutional	$filters = 6; kernel = 1; strides = 2$
2^{nd} Convolutional	$filters = 6; kernel = 3; strides = 2$
3^{rd} Convolutional	$filters = 6; kernel = 3; strides = 1$
4^{th} Convolutional	$filters = 6; kernel = 3; strides = 1$
5^{th} Convolutional	$filters = 6; kernel = 3; strides = 1$

TABLE 3: Residual Block Hyperparams

Layer	**Value**
Average Pooling	$strides = 2$
1^{st} Fully-Connected	$units = 20$
2^{st} Fully-Connected	$units = 10$
3^{rd} Fully-Connected (Output Layer)	$units = 1$

The model just defined is tested with every synthetic dataset of Table 1 concatenated with the real value of the BIDMC. The performance of the RR estimation using MAE with cross validation are presented in Table 4, with the mean and standard deviation of the evaluation obtained by the different folds. The best results are obtained with $AM = 0.2, BW = 0.2$ while the $General Amplitude$ gives the worst results.
$AM = 0.3, BW = 0.3$ gets closer to the best result, in particular the standard deviation is lower, so also this type of modulation can be considered as a good result.

TABLE 4: Residual Block Hyperparams

Layer	Value
$Standard(AM = 0.1, BW = 0.1, FM = 0.05$	4.55 ± 1.76
$AM = 0.2, BW = 0.2$	2.84 ± 0.54
$AM = 0.3, BW = 0.3$	3.09 ± 0.44
$AM = 0.4, BW = 0.4$	5.56 ± 5.67
$FM = 0.03$	8.39 ± 5.55
$FM = 0.025$	9.12 ± 5.09
$GeneralAmplitude = 3$	15.19 ± 0.0
$GeneralAmplitude = 6$	15.19 ± 0.0

VII. Concluding Remarks

The present work describes a Deep Learning approach to face a concrete and modern problem, the estimation of the Repiratory Rate. The framework proposed utilizes PPG signal and it's performance is maximized with respect to the data recorded by a professional pulse oximiter. This encourages a future development of a mobile application that provides the respiratory rate of a subject in real time. In the future some improvements could be done by varying the windows size chosen for the estimation or changing the architecture proposed with the latest improvement of Deep Learning models. Moreover some more powerful hardware could be employed in order to train bigger models with large amount of input data. What I have learned during this work is how Deep Learning can be used to resolve healthy problems and how the biometric signals can give many information that can be processed. The difficulties I faced are related to the development of the Deep Learning model, for its complexity and also for understanding the medical research done for the Respiration Rate.

VIII. Exam rules

References

[1] J. F. Fieselmann, M. S. Hendryx, C. M. Helms, and D. S. Wakefield, "Respiratory rate predicts cardiopulmonary arrest for internal medicine inpatients," *J Gen Intern Med*, vol. 8, pp. 354–360, Jul 1993.

[2] N. A. Chatterjee, P. N. Jensen, A. W. Harris, D. D. Nguyen, H. D. Huang, R. K. Cheng, J. J. Savla, T. R. Larsen, J. M. D. Gomez, J. M. Du-Fay-de Lavallaz, R. N. Lemaitre, B. McKnight, S. A. Gharib, and N. Sotoodehnia, "Admission respiratory status predicts mortality in COVID-19," *Influenza Other Respir Viruses*, vol. 15, pp. 569–572, 09 2021.

[3] D. Clifton, J. G. Douglas, P. S. Addison, and J. N. Watson, "Measurement of respiratory rate from the photoplethysmogram in chest clinic patients," *J Clin Monit Comput*, vol. 21, pp. 55–61, Feb 2007.

[4] D. Bian, P. Mehta, and N. Selvaraj, "Respiratory rate estimation using ppg: A deep learning approach," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*, pp. 5948–5952, 2020.

[5] P. H. Charlton, D. A. Birrenkott, T. Bonnici, M. A. F. Pimentel, A. E. W. Johnson, J. Alastruey, L. Tarassenko, P. J. Watkinson, R. Beale, and D. A. Clifton, "Breathing Rate Estimation From the Electrocardiogram and Photoplethysmogram: A Review," *IEEE Rev Biomed Eng*, vol. 11, pp. 2–20, 2018.

[6] C.-T. Yen, S.-N. Chang, and C.-H. Liao., "Deep learning algorithm evaluation of hypertension classification in less photoplethysmography signals conditions," *SAGE journals*, vol. 54, 03 2021.

[7] A. Tazarv and M. Levorato, "A deep learning approach to predict blood pressure from ppg signals," *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pp. 5658–5662, 2021.

[8] S. A. H. Aqajari, R. Cao, A. H. A. Zargari, and A.-M. Rahmani, "An end-to-end and accurate ppg-based respiratory rate estimation approach using cycle generative adversarial networks," *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pp. 744–747, 2021.

[9] M. N. I. Shuzan, M. H. Chowdhury, M. E. H. Chowdhury, M. M. Uddin, A. Khandakar, Z. B. Mahbub, and N. Nawaz, "A novel non-invasive estimation of respiration rate from photoplethysmograph signal using machine learning model," *ArXiv*, vol. abs/2102.09483, 2021.

[10] V. Ravichandran, B. Murugesan, V. Balakarthikeyan, K. Ram, P. Sp, J. Joseph, and M. Sivaprakasam, "Respnet: A deep learning model for extraction of respiration from photoplethysmogram," vol. 2019, pp. 5556–5559, 07 2019.

[11] A. SchÃ¤fer and K. W. Kratky, "Estimation of breathing rate from respiratory sinus arrhythmia: comparison of various methods," *Ann Biomed Eng*, vol. 36, pp. 476–485, Mar 2008.

[12] P. H. Charlton, T. Bonnici, L. Tarassenko, D. A. Clifton, R. Beale, and P. J. Watkinson, "An assessment of algorithms to estimate respiratory rate from the electrocardiogram and photoplethysmogram," *Physiological Measurement*, vol. 37, pp. 610–626, mar 2016.

[13] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, pp. E215–220, Jun 2000.

[14] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *CoRR*, vol. abs/1512.03385, 2015.