2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Feature Learning for Blood Pressure Estimation from Photoplethysmography

Clémentine Aguet, Jérôme Van Zaen, João Jorge, Martin Proença, Guillaume Bonnier, Pascal Frossard, Mathieu Lemay

Abstract— **Blood pressure (BP) is an important indicator for prevention and management of cardiovascular diseases. Alongside the improvement in sensors and wearables, photoplethysmography (PPG) appears to be a promising technology for continuous, non-invasive and cuffless BP monitoring. Previous attempts mainly focused on features extracted from the pulse morphology. In this paper, we propose to remove the feature engineering step and automatically generate features from an ensemble average (EA) PPG pulse and its derivatives, using convolutional neural network and a calibration measurement. We used the large VitalDB dataset to accurately evaluate the generalization capability of the proposed model. The model achieved mean errors of -0.24 ± 11.56 mmHg for SBP and -0.5 ± 6.52 mmHg for DBP. We observed a considerable reduction in error standard deviation of above 40% compared to the control case, which assumes no BP variation. Altogether, these results highlight the capability to model the dependency between PPG and BP.**

I. Introduction

Hypertension or persistently elevated blood pressure (BP) is a common life-threatening condition, affecting approximately one-third of the adult population. It is a key factor of cardiovascular diseases (CVDs), the leading cause of death worldwide [1]. Generally, without warning signs and symptoms, most people are unaware of the issue in its early stages. Moreover, BP can be rapidly affected by external factors, e.g., physical activity, emotions, and drugs. For these reasons, it is very important to monitor such a vital physiological parameter regularly and continuously. It makes an early diagnosis and, in turn, proper management of hypertension and related CVDs possible. The gold standard for continuous BP measurement is the arterial line, which has high risk of infection and is limited to a clinical environment. Therefore, sphygmomanometry is conventionally taken as reference for non-invasive BP monitoring. However, such cuff-based approach is uncomfortable for patients, and it is not suitable for continuous measurements.

To overcome these limitations, improving non-invasive, cuffless and continuous BP monitoring has been a major focus of research in recent years. With the growing presence of wearable, the use of a single biomedical signal for such a task has attracted a lot of attention. Photoplethysmography (PPG) is well-suited due to its simplicity, low cost and possible application to wearables and smartphones. This optical technique measures the blood volume variations in the microvascular bed of tissues using a light-emitting diode and a photodetector. Despite having a complex relationship, the PPG has similarity with the arterial BP morphology [2], showing its potential for BP estimation. Even if PPG is affected by local phenomena, its waveform results from the propagation of the central pressure pulse to the periphery and its morphology carries information about central BP. It is precisely this information we seek to extract to remain unaffected by local effects. Different approaches were investigated. Notably, methods based on pulse wave analysis (PWA) of the PPG waveform and its derivatives seem promising, and have revealed some time- and frequency-domain features that play a key role in the modeling of BP [3]–[5]. A regression model is then applied to map the features into BP values. Several machine learning (ML) methods have also been considered for this task, e.g., support vector regression [6], regression trees [7] and neural networks (NN) [4]. However, most of the current studies addressing BP estimation from PPG signals are mainly based on features derived from the pulse morphology. Such a feature engineering process highly depends on the signal quality. Furthermore, the computation of some complex features is time consuming and requires expert knowledge.

Alternatively, a model can automatically derive its own features – a process known as feature learning. The idea is to capture dynamic information of the PPG signal in a relevant latent space. It can be achieved with supervised NN. Although promising, the main limitations of such an approach are the large amount of data required to train the model and the possible loss of interpretability. Finding the best architecture and tuning the hyperparameters is also challenging. The majority of proposed models combine multiple physiological signals, typically PPG and electrocardiogram (ECG) [8]. To the best of our knowledge, only a few approaches attempted to take raw signals directly as input. The work in [9] proposed a complex spectro-temporal model based on PPG, its derivatives and spectrograms. More recently, the authors of [10] proposed to estimate BP from the spectrogram of a short PPG window and to use a Siamese architecture for calibration. Nevertheless, the results of these two attempts were insufficient for medical application. The authors of [11] estimated the BP from PPG signal and derivatives combining a convolutional NN (CNN) and a fully connected network. This study used a small dataset, making it difficult to evaluate the generalization capability.

In the context of BP monitoring from PPG signals, we have previously developed oBPM® [12]. This algorithm based on PWA has shown to accurately track systolic and mean BP

Corresponding e-mail: clementine.aguet@csem.ch

C. Aguet, J. Van Zaen, J. Jorge, M. Proença, G. Bonnier and M. Lemay are with the Swiss Center for Electronics and Microtechnology (CSEM), Neuchâtel. Switzerland

C. Aguet and P. Frossard are with the Signal Processing Laboratory (LTS4), Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

978-1-7281-1178-0/21/$31.00 ©2021 IEEE

changes in patients undergoing general anesthesia [13]. In parallel, we studied the potential of a data-driven approach, by integrating ML techniques into oBPM® technology [14]. In the present work, we investigate a feature learning approach for such a task. The model is based on an ensemble averaged pulse (EA) computed over a PPG window, without any requirements of feature engineering. It combines a CNN for feature extraction and a fully connected network to estimate BP from these features. Taking advantage of the large open database VitalDB [15], our model was train and evaluated on data with acute BP variations. The calibration is done by incorporating an initial measurement in the model. Such a procedure is well appropriate for a possible use case.

II. Materials and Methods

A. Dataset

The proposed method was trained and tested on part of the data retrieved from the VitalDB data bank. This open database was collected at the Seoul National University Hospital department of Anesthesia (Seoul, Republic of Korea). The study was approved by the local ethics committee (H-1408-101-605, NCT02914444 at ClinicalTrials.gov). The experimental setup is described in [15]. The database includes multiple biosignals of adult patients recorded during various non-cardiac surgeries (lasting in average 3 hours and going up to more than 16 hours). We selected those including arterial line and PPG signals, both necessary for our research, which reduced the database to 3326 patients. Interestingly, due to drug administration, this dataset includes large BP variations over time, providing a wide range of BP values over which to train and test our approach.

B. Pulse ensemble averaging

After filtering the raw PPG signal to remove the baseline and reduce noise, the signal was then split into non-overlapping 20-second windows and aggregated into ensemble-averaged (EA) pulses. More specifically, one EA pulse, its first (velocity plethysmography: VPG) and second (acceleration plethysmography: APG) derivatives, its cardiac period, and an associated quality index was extracted from each window. This process is described in more detail in [13]. The quality index (≥ 80%) and the reference variability (standard deviation/average ≤ 10%) were then used to exclude windows with too much noise or signal distortions. Patients with less than 5 EA pulses were also excluded. To standardize their lengths, the EA pulses were padded with zeros up to the maximum observed cardiac period and then resampled to a length of 256. The overall process resulted in 1567 patients and a total of 126'327 EA pulses. Some BP characteristics of this dataset are presented in TABLE I. To ensure reliability of the results, 80% of the data was used to train the model, while the remaining 20% was kept for evaluation. Patients were carefully distributed across the two sets in a stratified manner, using the mean and standard deviation of the systolic BP (SBP). Each patient only contributed to the train or test set.

TABLE I. Dataset Characteristics (N = 1567)

Characteristics	Mean ± STD	(Range)
SBP (mmHg)	115.4 ± 13.7	(50.8 – 260.5)
DBP (mmHg)	61.4 ± 7.4	(30.1 – 143.5)
MAP (mmHg)	81.1 ± 9.9	(37.2 – 196.9)

C. Calibration approach

An initial calibration process is usually necessary to correct potential baseline drift. Therefore, we selected an approach including one initial measurement. The model took as inputs a calibration EA pulse with the corresponding BP value and an estimation EA pulse. The calibration pulse was always taken from the same recording as, and before the occurrence of, the estimation EA pulse. The average timespan between calibration EA pulse and estimation EA pulse (Δt) is 74 minutes, ranging from 1 minutes to over 16 hours.

D. Model architecture

A schematic description of the overall model is illustrated in Figure 1. The two feature extractors were identical, sharing the same architecture and parameters. They received in parallel the estimation and calibration inputs. The output features were then concatenated together with the BP value of the calibration segment. The BP values of the estimation segment were used as ground truth during the training process. The resulting feature vector was finally fed into a regression model for the BP estimation.

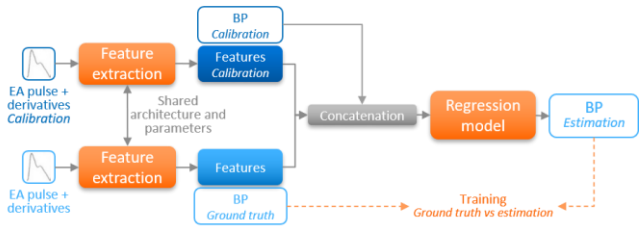

Figure 1. Learning approach with a calibration measure.

We proposed to take advantage of CNN to extract features from temporal domain inputs, including the EA pulse and its derivatives. The CNN architecture is presented in Figure 2. In total, 4 convolutional blocks were stacked to extract temporal features. A rectified linear unit (ReLU) activation function was applied after each 1-dimensional convolution. The last layer of each block was a max-pooling layer to reduce the inputs dimensionality. Finally, a dropout layer came after the 4 convolutional blocks to reduce overfitting. The concatenated features were then fed into a series of two fully connected (FC) layers with ReLU activation functions. And the final ReLU outputted the estimated BP value. The network weights were initialized using the He method proposed in [16], which is well-suited to ReLU like activations.

Figure 2. CNN architecture for feature extraction. The information in brackets indicates to the number of channels and the kernel size.

E. Training setting

The model was implemented in Python using the PyTorch deep learning framework. Two models were trained in a fully supervised manner. One to output the systolic BP value (SBP) and the other the diastolic BP value (DBP). The Adam algorithm was chosen to optimize the network parameters with $\beta_1 = 0.9$, $\beta_2 = 0.999$ and a weight decay of 10^{-4}. The learning rate was set to 10^{-4}. The BP estimation being a

regression problem, the Huber loss was selected as the loss function. This smooth version of the L1 loss is more robust to outliers than the mean square error (MSE) and more efficient than the mean absolute error (MAE) due to continuous derivative. The model was trained over 25 epochs with a batch size set to 32. All hyperparameters were initially defined based on a literature review and then experimentally adapted using 10-fold cross-validation on the training data.

F. Evaluation metrics

The performance was assessed by comparing the estimates of the proposed model to their corresponding invasive reference values. By analogy with the ISO 81060-2:2018 norm [17], not fully applicable to noninvasive cuffless methods, we computed the mean error (ME) between the estimated and reference BP values along with the standard deviation of this error (STDE). A flat model was used as baseline performance. With the assumption of no BP change, it outputs the calibration reference as BP estimate. Beyond simply being a control case, its STDE also gives interesting insights regarding the BP variations in the dataset.

III. Results

TABLE II summarizes the overall performance of the proposed model with different input combinations for SBP estimation in terms of ME and STDE in mmHg. The best performance was obtained when using the PPG and the APG as inputs. This model achieved an estimation error of -0.24 ± 11.56 mmHg for SBP. The performance for DBP is presented in TABLE *III*, the estimation error was -0.50 ± 6.52 mmHg.

TABLE II. SBP ESTIMATION PERFORMANCES IN MMHG

	Train		Test	
Model	ME	STDE	ME	STDE
PPG	0.28	9.43	0.41	11.58
PPG + APG	**-0.31**	**10.27**	**-0.24**	**11.56**
PPG + VPG + APG	-0.65	9.60	-0.46	11.75
Flat model	0.08	20.59	0.02	21.05

TABLE III. DBP ESTIMATION PERFORMANCES IN MMHG

	Train		Test	
Model	ME	STDE	ME	STDE
PPG	-0.25	5.25	-0.33	6.65
PPG + APG	**-0.40**	**5.62**	**-0.50**	**6.52**
PPG + VPG + APG	0.42	5.09	0.43	6.60
Flat model	0.10	11.09	-0.02	11.34

The 2D histograms plots in Figure 3 and Figure 4 illustrate the correlation between estimated BP values and the corresponding arterial BP value in the test set for SBP and DBP, respectively. The correlation coefficient is 0.84 for SBP and 0.85 for DBP.

IV. Discussion

To the best of our knowledge, few studies have documented the usefulness of convolutional networks for cuffless BP estimation, even less exploiting PPG without adding concurrent ECG signals. In this paper, we proposed a novel BP estimation method based on CNN for automatic feature extraction that achieves promising results.

The first question concerned the choice of starting signals. Unlike most proposed methods, the idea here was to only use a PPG sensor, which is more suitable for long-term wearable

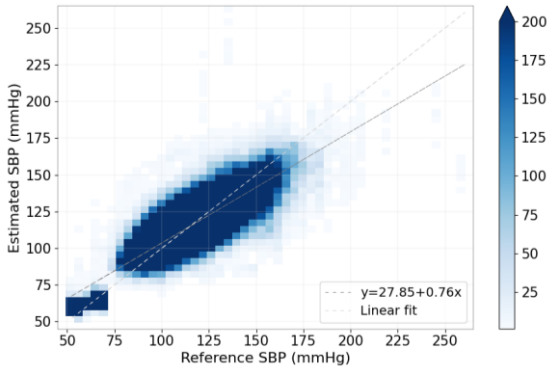

Figure 3. 2D histogram of estimated and reference SBP in mmHg with a linear fit (dark gray) and perfect correlation (light gray).

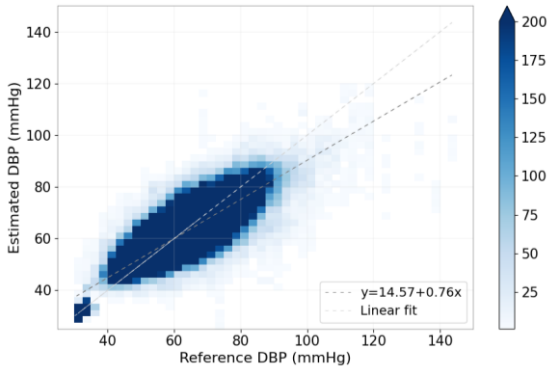

Figure 4. 2D histogram of estimated and reference DBP in mmHg with a linear fit (dark gray) and perfect correlation (light gray).

applications. Nevertheless, combining the PPG pulse with its second derivative tended to yield more generalizable features and therefore to reduce the overfitting between the training and test performance metrics (TABLE *II*). This observation is not surprising and is consistent with the literature review and feature engineering approaches, where characteristic points of the APG are typically used in BP estimation model [5]. It confirms that APG carries more relevant information about arterial stiffness and indirectly about BP than VPG [18].

The proposed model significantly outperformed the flat model, with a reduction in STDE of approximately 45% for SBP and 43% for DBP, thereby confirming that the PPG and its derivatives enclose information relating to the BP. The estimation error was -0.24 ± 11.56 mmHg for SBP and -0.50 ± 6.52 mmHg for DBP. Such ranges when dealing with large-scale data highlight the difficulty of developing a robust and generalizable model. The higher precision for DBP estimation compared to SBP estimation might be partly explained by the lower variance of DBP (flat model in TABLE *II* vs TABLE *III*). A possible limitation of the present architecture is that two separate models were trained to output SBP and DBP, respectively. As those factors are closely related, a single model could be trained to simultaneously estimate SBP and DBP. With such multitask learning approach, the model could better capture shared representation and avoid overfitting. This idea will be investigated in future work. The 2D histograms for SBP and DBP in Figure 3 and Figure 4 show that the proposed model tends to underestimate high BP and overestimate low BP. The reason behind this is unclear and needs further investigation.

A direct comparison of the results with the literature is challenging, as multiple factors alter the estimation error, including the pool of subjects, the degree of BP variations or the calibration method. As far as we know, only [19] presented a PPG-based BP estimation model trained and evaluated on this large VitalDB database. Although their model obtained better performances (0.016 ± 7.66 mmHg for SBP and -0.043 ± 4.22 mmHg for DBP), important aspects should be highlighted. This method was based on PPG morphology features, whereas our approach removed the feature engineering step and achieved automatic feature extraction. Furthermore, the difference in performance is mainly due to the impact of calibration method. In [19], the offset was adjusted using patient's recording average BP values. Such a procedure is appropriate for a feasibility study but hardly applicable in practice. Our proposed model was designed to include one initial measurement. It is more representative of a possible use case, where the calibration is typically done using a cuff measurement at the doctor's office. The necessity of a calibration process is one of the main limitations of current attempts for cuffless BP estimation model. The need of a measure with an existing acceptable standard might increase the complexity of potential applications. An alternative could be to simplify the calibration by using the patient's personal information, such as age, weight, height, and gender. A deeper study should therefore be conducted to assess the dependence on calibration, and its frequency, and therefore evaluate the stability of the approach.

Besides the large number of patients needed for NN training, the dataset used has other interesting aspects. Recorded during surgical procedures with anesthetic drug administration, it includes more acute BP variations over time than most available datasets. Despite the promising results, a more in-depth evaluation of the model ability to track these acute BP changes relative to small variations should be conducted.

One challenge faced by NN is the high dependence on the hyperparameter configuration. Some hyperparameters are related to the network architecture (number of layers and units), while others to the training algorithm (learning rate, number of epochs, batch size). As previously mentioned, they were here initialized based on a literature review and experimentally adjusted using cross-validation in the training set. However, the performance could be further improved by tuning those hyperparameters further. Bayesian optimization appears to be a promising approach for this task.

V. Conclusion

The results obtained in this study highlighted the potential of feature learning approaches with CNNs for cuffless BP estimation exclusively based on a PPG sensor. Such an approach seems to benefit from the PPG second derivative to improve its generalization capability. Furthermore, the chosen calibration process of using a simple initial measure brings it one step closer to an actual application. The proposed model was trained and evaluated on a large dataset, which revealed its ability to track a wide range of BP variations. The results are promising and motivate further investigations.

References

[1] K. T. Mills, A. Stefanescu, and J. He, "The global epidemiology of hypertension," *Nat. Rev. Nephrol.*, vol. 16, no. 4, Apr. 2020.

[2] G. Martínez, N. Howard, D. Abbott, K. Lim, R. Ward, and M. Elgendi, "Can Photoplethysmography Replace Arterial Blood Pressure in the Assessment of Blood Pressure?," *J. Clin. Med.*, vol. 7, Sep. 2018.

[3] X. F. Teng and Y. T. Zhang, "Continuous and noninvasive estimation of arterial blood pressure using a photoplethysmographic approach," in *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE Cat. No.03CH37439)*, Cancun, Mexico, 2003, pp. 3153–3156.

[4] Y. Kurylyak, F. Lamonaca, and D. Grimaldi, "A Neural Network-based method for continuous blood pressure estimation from a PPG signal," in *2013 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, Minneapolis, MN, USA, May 2013.

[5] M. Liu, L.-M. Po, and H. Fu, "Cuffless Blood Pressure Estimation Based on Photoplethysmography Signal and Its Second Derivative," *Int. J. Comput. Theory Eng.*, vol. 9, no. 3, pp. 202–206, 2017.

[6] Y. Zhang and Z. Feng, "A SVM Method for Continuous Blood Pressure Estimation from a PPG Signal," in *Proceedings of the 9th International Conference on Machine Learning and Computing - ICMLC 2017*, Singapore, Singapore, 2017, pp. 128–132.

[7] S. G. Khalid, J. Zhang, F. Chen, and D. Zheng, "Blood Pressure Estimation Using Photoplethysmography Only: Comparison between Different Machine Learning Approaches," *J. Healthc. Eng.*, vol. 2018, pp. 1–13, Oct. 2018.

[8] S. Baek, J. Jang, and S. Yoon, "End-to-End Blood Pressure Prediction via Fully Convolutional Networks," *IEEE Access*, 2019.

[9] G. Slapničar, N. Mlakar, and M. Luštrek, "Blood Pressure Estimation from Photoplethysmogram Using a Spectro-Temporal Deep Neural Network," *Sensors*, vol. 19, no. 15, p. 3420, Aug. 2019.

[10] O. Schlesinger, N. Vigderhouse, D. Eytan, and Y. Moshe, "Blood Pressure Estimation From PPG Signals Using Convolutional Neural Networks And Siamese Network," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Barcelona, Spain, May 2020, pp. 1135–1139.

[11] S. Shimazaki, H. Kawanaka, H. Ishikawa, K. Inoue, and K. Oguri, "Cuffless Blood Pressure Estimation from only the Waveform of Photoplethysmography using CNN," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Berlin, Germany, Jul. 2019, pp. 5042–5045.

[12] M. Proença, J. Sola, M. Lemay, and C. Verjus, "Method apparaus and computer program for determining a blood pressure value," Jun. 18, 2017.

[13] Y. Ghamri *et al.*, "Automated Pulse Oximeter Waveform Analysis to Track Changes in Blood Pressure During Anesthesia Induction: A Proof-of-Concept Study," *Anesth. Analg.*, vol. 130, no. 5, pp. 1222–1233, May 2020.

[14] J. Jorge *et al.*, "Machine Learning Approaches For Improved Continuous, Non-occlusive Arterial Pressure Monitoring Using Photoplethysmography," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Montreal, QC, Canada, Jul. 2020, pp. 910–913.

[15] H.-C. Lee and C.-W. Jung, "Vital Recorder—a free research tool for automatic recording of high-resolution time-synchronised physiological data from multiple anaesthesia devices," *Sci. Rep.*, vol. 8, no. 1, p. 1527, Dec. 2018.

[16] K. He, X. Zhang, S. Ren, and J. Sun, "Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification," *ArXiv150201852 Cs*, Feb. 2015, Accessed: Jan. 29, 2021. [Online].

[17] International Organization for Standardization (ISO), "ISO 81060-2:2018 Non-invasive sphygmomanometers — Part 2: clinical investigation of intermittent automated measurement type." 2018.

[18] M. Proença *et al.*, "Pulse Wave Analysis Techniques," in *The Handbook of Cuffless Blood Pressure Monitoring: A Practical Guide for Clinicians, Researchers, and Engineers*, J. Solà and R. Delgado-Gonzalo, Eds. Cham: Springer International Publishing, 2019, pp. 107–137.

[19] S. Yang, J. Sohn, S. Lee, J. Lee, and H. C. Kim, "Estimation and Validation of Arterial Blood Pressure using Photoplethysmogram Morphology Features in conjunction with Pulse Arrival Time in Large Open Databases," *IEEE J. Biomed. Health Inform.*, pp. 1–1, 2020.