Hypertension Risk Assessment from Photoplethysmographic Recordings Using Deep Learning Classificators

Jesús Cano[1],Vicente Bertomeu-González[2], Lorenzo Fácia[3],
Roberto Zangróniz[4], Raúl Alcaraz[4], José J Rieta[1*]

[1] BioMIT.org, Electronic Engineering Department, Universitat Politecnica de Valencia, Spain
[2] Clinical Medicine Department, Miguel Hernández University, Elche Spain
[3] Cardiology Department, Hospital General Universitario de Valencia, Spain
[4] Research Group in Electronic, Biomed. and Telecomm. Eng., Univ. of Castilla-La Mancha, Spain

Abstract

Regular monitoring of blood pressure (BP) is essential to make an early detection of cardiovascular diseases caused by hypertension, a potentially deadly condition that do not present symptoms in its first stages. This study aims to investigate whether deep learning techniques can assess risk levels of BP using only photoplethysmographic (PPG) recordings without the need of electrocardiographic (ECG) recordings, as in many previous studies. 15.240 segments from 50 different patients containing simultaneous PPG and arterial blood pressure (ABP) signals were analysed. GoogleNet and ResNet pretrained convolutional neural networks (CNN) with the scalogram of PPG signals obtained by continuous wavelet transform (CWT) used as input images were employed for the classification. The highest F1 score was achieved by discriminating normotensive (NT) patients from prehypertensive (PH) and hypertensive (HT), being 92.10% *for GoogleNet and* 93.91% *for ResNet, respectively. In addition, intra-patient classification using different data segments for training and validation provided an F1 score of* 90.28% *with GoogleNet and* 89.04% *with ResNet. Time frequency transformation of PPG recordings to feed deep learning classifiers has been able to provide outstanding results in hypertension risk assessment without requiring either ECG recordings or feature extraction*

1. Introduction

Blood pressure (BP) is the most important biomarker for cardiovascular diseases, which is the leading cause of mortality worldwide and a major contributor to the reduction of quality of life [1]. Thus, early diagnosis and control of hypertension is essential for prevention. Moreover, most patients with severely elevated blood pressure have asymptomatic hypertension without signs or symptoms of end-organ damage. Without an early detection, it can derive in an hypertensive urgency with the presence of risk factors for progressive diseases as hearth failure and preexisting renal insufficiency or severe uncontrolled hypertension [2].

The most common noninvasive technique for BP measurement are based on uncomfortable arterial occlusion by inflatable cuffs where adequate accuracy is offered but only provide intermittent measurement and needs to be applied by professionals [3]. Recent advances in sensor technology have developed unobstructive cuffless devices to measure physiological parameters anytime. In this way, the use of photoplethysmographic (PPG) recordings is very promising for being noninvasive, with continuous measurement, low cost, simple and with a high correlation with arterial BP in frequency and time domain [4].

In this regard, many studies have applied artificial intelligence technology in order to estimate or discriminate between blood pressure levels. Machine learning techniques combining electrocardiographic (ECG) and PPG signals get use of propagation theory with parameters as pulse transit time (PTT), pulse arrival time (PAT) and pulse wave velocity (PWV) to determine cardiovascular state [5]. Recent studies combine this propagation parameters with PPG morphological feature extraction as inputs for the models [6]. In recent years, deep learning with powerful computational methods that eliminates feature extraction have shown an improvement in BP estimation from PPG signals [7]

The present work proposes a method for hypertension risk assessment using the pretrained CNNs GoogleNet and ResNet and the scalogram of PPG signals by continuous wavelet transform (CWT) as inputs to feed this models without the need for ECG recordings or hand extraction of signals features for classification.

2. Material and Methods

2.1. Data acquisition and preprocessing

A total number of 635 recordings from 50 different patients were collected from MIMIC-III ('Medical Information Mart for Intensive Care'), a free to use database with information and vital signs of patients admitted to critical care unit [8]. As signals are obtained by commercial devices, often contain artefacts caused by sensor movements or loss of contact. The main artefacts detected and the reason why many records were not incorporated in the dataset was noisy PPG and BP signals, improbable BP values, or continuous signals that do not represent their characteristic morphology.

In this study, simultaneous and stable ABP and PPG signals with 120 s length and 125 Hz sampling rates were employed. The systolic blood pressure (SBP) were extracted from the ABP signals, to which no preprocessing was carried out, and were used to label the PPG signals in normotensive (NT), prehypertensive (PH) and hypertensive (HT), with SBP values of 120 mmHg or lower, between 120 and 140 mmHg and 140 mmHg or higher respectively, as defined by the US National Institutes of Health [1]. Additionally, a 0.5-10 Hz Chebyshev II bandpass filter of fourth order was applied to the normalized PPG signals to remove noise [9]. After these steps, both signals were cut in 5s length segments, being 15.240 the total number of segments analysed.

Furthermore, it is studied the effect of down sampling the PPG signal to 25 Hz in order to introduce this method in the mobile health field. Lower sampling rates reduces the power consumption of the device and if the transmitted data is reduced, the database storage will be saved [10].

2.2. Pretrained Convolutional Neural Networks

Two different deep convolutional neural networks were evaluated for the hypertension risk classification problem, GoogleNet [11] and ResNet50 [12]. Transfer learning works with pretrained networks that have already used the ImageNet dataset [13] with more than a million of images and 1000 classes to learn how to extract informative features. They are a starting point for identification and classification problems much faster than training from scratch using less images for training. In this way, the last learning layer and the final layer of classification are replaced with new layers adjusted to the new training images.

Training options established a minimum Batch size of 128, the validation frequency is modified depending on the number of training images and the maximum Epoch was 25. In order to minimise the effect of overfitting, early stopping technique stops training automatically when the loss on the validation set starts to increase.

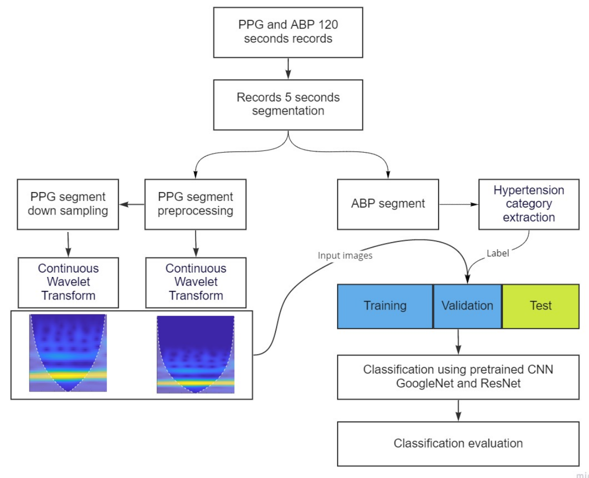

Figure 1. Deep learning classification process.

Due to both networks uses 224x224x3 sizes RGB images as inputs, PPG segments were processed by continuous wavelet transform and transformed to a scalogram, a representation of frequency along the time, and then resized to 224x224x3 to feed the training models. It has been used the analytic Morse (3,60) wavelet in the cwtfilterbank of Matlab, setting the VoicesperOctave to 12 to create the CWT. In the scalogram is included the cone of influence, that represents where occur edge effects in the CWT for obtaining better classification results.

2.3. Hypertension Risk Assessment

In order to classify the hypertension risk using PPG signals, the three group of labelled signals were compared discriminating NT versus PH + HT and NT + PH versus HT for the classification. Then, two different methods were developed, starting with the division of the dataset of 635 records in 80% for training and validation and 20% for testing with recordings of new patients. In addition, the first group is randomly divided again in 80% for training and 20% for validation. It must be taken into account that in this approach this division is done to the images, not to the records, so images obtained from segments of 5 seconds from the same records are present in training and validation datasets. Furthermore, intra-patient classification was developed using different PPG complete signals of 120 s for training and validation.

Finally, the evaluation of the results of validation was carried out by the statistical tests of sensibility (Se), specificity (SP) and F1 score, calculated from the recall and precision of the validation.

3. Results

Tables 1 and 2 shows the classification efficiency when the dataset is randomly divided in training and validation with the original sampling rate of 125 Hz and downsampled to 25 Hz respectively.

		Se	Sp	F1score
GoogleNet	NTvsPH+HT	95.44%	83.08%	92.10%
	NT+PHvsHT	84.45%	90.99%	84.16%
ResNet	NTvsPH+HT	94.37%	90.52%	93.91%
	NT+PHvsHT	87.67%	93.87%	88.23%

Table 1. Training-Validation performance from PPG signals with 125 Hz sampling rate

		Se	Sp	F1score
GoogleNet	NTvsPH+HT	94%	83.83%	91.54%
	NT+PHvsHT	82.49%	91.87%	83.64%
ResNet	NTvsPH+HT	94.56%	89.29%	93.58%
	NT+PHvsHT	88.94%	92.83%	88.08%

Table 2. Training-Validation performance from PPG signals with 25 Hz sampling rate

All the models followed the same pattern, in which F1 score percentage is higher comparing NT segments versus PH + HT, being over 91.5%, than comparing NT + PH segments versus HT, being lower than 88.5%.

Tables 3 and 4 show the test performance when the dataset is randomly divided in training (80%) and validation (20%) with the original sampling rate of 125 Hz and downsampled to 25 Hz respectively.

		Se	Sp	F1score
GoogleNet	NT vs PH+HT	68.13%	48.03%	65.95%
	NT+PH vs HT	9.05%	87.67%	10.07%
ResNet	NT vs PH+HT	69.67%	49.58%	67.31%
	NT+PH vs HT	16.11%	83.70%	15.38%

Table 3. Test results using GoogleNet and ResNet models from 125 Hz sampling rate PPG signals

		Se	Sp	F1score
GoogleNet	NT vs PH+HT	59.70%	51.16%	60.97%
	NT+PH vs HT	4.61%	89.47%	5.60%
ResNet	NT vs PH+HT	67.75%	45.68%	65.17%
	NT+PH vs HT	20.18%	82.33%	18.29%

Table 4. Test results using GoogleNet and ResNet models from 25 Hz sampling rate subsampled PPG signals

All models comparing NT vs PH+HT obtain similar F1 score, around 65%. Nevertheless, models comparing NT+PH vs HT obtain low test results, with F1 score around 10% as almost all new records are classified as negative.

After obtaining this results, the intra-patient study was developed only with the first distribution of hypertension risks as represented in Table 5. The higher F1 score percentage, approximately 90%, is obtained in models that uses subsampled signals representations as inputs.

		Se	Sp	F1score
GoogleNet	125 Hz	88.89%	78.13%	87.37%
	25 Hz	90.28%	85.42%	90.28%
ResNet	125 Hz	83.83%	83.83%	85.71%
	25 Hz	93.06%	76.04%	89.04%

Table 5. Intra patient classification performance from PPG signals comparing NT segments versus PH + HT

4. Discussion

Being able to monitor and detect hypertension with a continuous measurement is of great importance as is the main risk factor of many cardiovascular diseases. For this reason, new cuff-less devices based on Machine Learning and Deep Learning techniques have been proposed as an alternative to traditional methods to measure BP. Almost all the studies use the PPG signal as its variation in morphology is mainly caused by the activity of the heart and the condition of the vascular walls. Moreover, PPG signals are simple to obtain with nonivasive low cost devices and can be measured in real time.

This work tries to investigate whether new Deep Learning techniques are able to impose to the most commonly used technique, Machine Learning classifiers, to obtain the hypertension risk. This method has some difficulties, needing to extract features from signals with different properties and quality and in this particular type of classification, the need to use other biomedical signal, the ECG, to obtain PAT. Deep learning approaches is a good solution to avoid this difficulties as extracts the features from an image obtained from PPG signals (without needing ECG signal) automatically and robustly.

F1 score for both pretrained models and using both the original PPG signal with 125 Hz sampling rate and subsampled to 25 Hz have been around 93%. This similar and huge results between both sampling rates lead to the proposal to use subsampled signals for the potential application of this technique in a wearable device as would reduce the computational complexity and use of memory.

It is noteworthy that in all models, the best results are obtained comparing NT with PH and HT patients. This is very relevant, since the best classification performance is obtained when PH patients are identified as diseased, so the extracted features in PH patients are more similar to those in HT patients than in NT patients. In addition, it should be taken into account that PH patients do not show serious symptoms until they are in very advanced stages of the disease, causing serious cardiovascular problems, so alerting this group as sick patients is very interesting.

Furthermore, in test results, the classification of new subjects is not correctly performed, with F1 score around 65%. This could be because each subject have a unique cardiovascular dynamics, so not everyone has the same relationship between PPG signals and BP. Moreover, the same PPG cycle shapes do not guarantee the same BP level, so information from each subject should be in training, validation and test sets [14].

Finally, intra-patient classification has been studied, where segments of PPG signal from the same subject in different hours have been introduced in training and validation sets seeking to avoid the possible overfitting. GoogleNet and subsampled PPG signals has obtained the higher F1 score value, 90.28%, proving the importance of calibration to reduce the classification error.

5. Conclusion

Hypertension risk classification models from PPG recordings using deep learning classifiers have been evaluated. The combination of continuous wavelet transform and pretrained CNN models has demonstrated outstanding performance and potential as extracts the main features from PPG image representation automatically and does not need ECG signals for the feature extraction as classical methods based on Machine Learning, with F1 score of 93.91%. In addition, the viability of use subsampled PPG signals and intrapatient accomplishment has been proved for the implementation of this models in wearable devices.

Acknowledgements

Research supported by grants DPI2017–83952–C3 from MINECO/AEI/FEDER UE and SBPLY/17/180501/000411 from JCCLM.

References

[1] Chobanian AV, Bakris GL, Black HR, Cushman WC, Green LA, Izzo JL, Jones DW, Materson BJ, Oparil S, Wright JT, Roccella EJ. Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure. Hypertension dec 2003;42(6):1206–1252. ISSN 0194-911X.

[2] Kessler CS, Joudeh Y. Evaluation and Treatment of Severe Asymptomatic Hypertension. American Family Physician feb 2010;81(4):470–476. ISSN 0002-838X.

[3] Pielmus AG, Mühlstef J, Bresch E, Glos M, Jungen C, Mieke S, Orglmeister R, Schulze A, Stender B, Voigt V, Zaunseder S. Surrogate based continuous noninvasive blood pressure measurement. Biomedizinische Technik feb 2021; ISSN 00135585.

[4] Martínez G, Howard N, Abbott D, Lim K, Ward R, Elgendi M. Can Photoplethysmography Replace Arterial Blood Pressure in the Assessment of Blood Pressure? Journal of Clinical Medicine sep 2018;7(10):316. ISSN 2077-0383.

[5] Yoon YZ, Kang JM, Kwon Y, Park S, Noh S, Kim Y, Park J, Hwang SW. Cuff-Less Blood Pressure Estimation Using Pulse Waveform Analysis and Pulse Arrival Time. IEEE Journal of Biomedical and Health Informatics jul 2018; 22(4):1068–1074. ISSN 2168-2194.

[6] Kachuee M, Kiani MM, Mohammadzade H, Shabany M. Cuffless Blood Pressure Estimation Algorithms for Continuous Health-Care Monitoring. IEEE Transactions on Biomedical Engineering apr 2017;64(4):859–869. ISSN 0018-9294.

[7] Liang Y, Chen Z, Ward R, Elgendi M. Photoplethysmography and deep learning: Enhancing hypertension risk stratification. Biosensors oct 2018;8(4). ISSN 20796374.

[8] Johnson AE, Pollard TJ, Shen L, Lehman LWH, Feng M, Ghassemi M, Moody B, Szolovits P, Anthony Celi L, Mark RG. MIMIC-III, a freely accessible critical care database. Scientific Data may 2016;3(1):1–9. ISSN 20524463.

[9] Liang Y, Elgendi M, Chen Z, Ward R. Analysis: An optimal filter for short photoplethysmogram signals. Scientific Data may 2018;5(1):1–12. ISSN 20524463.

[10] Fujita D, Suzuki A. Evaluation of the Possible Use of PPG Waveform Features Measured at Low Sampling Rate. IEEE Access 2019;7:58361–58367. ISSN 21693536.

[11] Szegedy C, Wei Liu, Yangqing Jia, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. In 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), volume 07-12-June. IEEE. ISBN 978-1-4673-6964-0. ISSN 10636919, jun 2015; 1–9.

[12] He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), volume 2016-Decem. IEEE. ISBN 978-1-4673-8851-1. ISSN 10636919, jun 2016; 770–778.

[13] Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M, Berg AC, Fei-Fei L. ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision dec 2015; 115(3):211–252. ISSN 0920-5691.

[14] Esmaelpoor J, Moradi MH, Kadkhodamohammadi A. A multistage deep neural network model for blood pressure estimation using photoplethysmogram signals. Computers in Biology and Medicine may 2020;120:103719. ISSN 00104825.

Address for correspondence:

José J. Rieta

BioMIT.org, Electronic Engineering Department, Building 7F-5

Universitat Politecnica de Valencia

Camino de Vera, s/n, 46022, Valencia, Spain

E-mail: jjrieta@upv.es