BLOOD PRESSURE ESTIMATION FROM PPG SIGNALS USING CONVOLUTIONAL NEURAL NETWORKS AND SIAMESE NETWORK

Oded Schlesinger[*1], *Nitai Vigderhouse*[*1], *Danny Eytan*[†], *Yair Moshe*[*]

[*]Signal and Image Processing Laboratory (SIPL)
Andrew and Erna Viterbi Faculty of Electrical Engineering
Technion – Israel Institute of Technology
http://sipl.technion.ac.il/

[†]Ruth and Bruce Rappaport Faculty of Medicine
Technion – Israel Institute of Technology

ABSTRACT

Blood pressure (BP) is a vital sign of the human body and an important parameter for early detection of cardiovascular diseases. It is usually measured using cuff-based devices or monitored invasively in critically-ill patients. This paper presents two techniques that enable continuous and noninvasive cuff-less BP estimation using photoplethysmography (PPG) signals with Convolutional Neural Networks (CNNs). The first technique is calibration-free. The second technique achieves a more accurate measurement by estimating BP changes with respect to a patient's PPG and ground truth BP values at calibration time. For this purpose, it uses Siamese network architecture. When trained and tested on the MIMIC-II database, it achieves mean absolute difference in the systolic and diastolic BP of 5.95 mmHg and 3.41 mmHg respectively. These results almost comply with the AAMI recommendation and are as accurate as the values estimated by many home BP measuring devices.

Index Terms— Blood pressure, convolutional neural network (CNN), noninvasive, photoplethysmography (PPG), Siamese network

1. INTRODUCTION

Blood pressure (BP) is the result of force exerted in the arteries by blood as it circulates. It is usually expressed in terms of systolic pressure (when the heart beats and BP is at its highest) and diastolic pressure (between heart beats, when BP is at its lowest) and measured in millimeters of mercury (mmHg). Normal resting BP in an adult is approximately 120 mmHg systolic and 80 mmHg diastolic. BP is an important parameter of the human body whose measurement allows for the early detection of medical issues, especially cardiovascular diseases, which are a leading cause of mortality and morbidity worldwide. High BP, hypertension, is a major risk for dangerous health conditions such as stroke or heart attack. Low BP, hypotension, can cause dizziness and fainting or may indicate serious heart, endocrine or neurological disorders. Therefore, it is highly important to measure BP routinely. Continuous monitoring of BP, along with monitoring of other vital signs, allows an accurate evaluation of the patient's physiological state, prompt detection of deteriorations and their prediction. The current widespread BP monitoring methods are divided into invasive and noninvasive methods. Invasive arterial line is a clinical standard for continuous high accuracy BP measurement. However, it has adverse effects associated with invasive measurements, such as potential infection, all of them are associated with an increased morbidity. Noninvasive BP measurement methods typically use an oscillometry inflatable arm or wrist cuff. These methods are not feasible for long-term ambulatory BP monitoring due to discomfort caused by repeated inflation and deflation and mobility limitations caused by the measuring device.

The Association for the Advancement of Medical Instrumentation (AAMI) recommends that the mean absolute difference (MAD) of noninvasive BP measurement technologies should not be greater than 5 mmHg and the standard deviation should not be greater than 8 mmHg compared to a reference method [1]. Home BP measuring devices may be inaccurate in 5% to 15% of patients, and a difference of 5 mmHg or higher is common [2, 3]. Therefore, BP measurements of other noninvasive technologies are expected to have at least similar accuracy, so they are at least as reliable as home BP measuring devices.

Photoplethysmography (PPG) is an optically obtained signal that can be used to detect blood volume changes in the microvascular bed of tissue. It is obtained by illuminating the skin and measuring changes in light absorption. In a clinical setting, this signal is often obtained by a pulse oximeter while in an ambulatory setting, this signal can be obtained by a smartwatch or other mobile devices. PPG signals contain information on cardiovascular parameters such as heart rate, blood oxygen saturation and BP. Since PPG is noninvasive, simple and low-cost, it has wide potential for clinical

[1] The first two authors contributed equally to this work.

applications [4]. This is a very promising technology and a potentially feasible replacement for invasive arterial BP measurement, as PPG has been shown to have a high correlation with arterial BP in the time and frequency domain [5]. However, PPG signals are prone to significant artifacts that adversely affect measurement accuracy [6]. It is therefore highly important to evaluate the quality of these signals when used for medical applications.

Many attempts have been made to estimate BP using PPG signals. These attempts can be divided into three approaches – pulse transit time (PTT), pulse wave velocity (PWV) and pulse wave analysis (PWA) [7]. PTT uses PPG and electrocardiogram (ECG) to calculate the time required for the BP to move between two points on an arterial line. PWV uses two PPG signals obtained at a known distance apart along the same arterial branch to calculate the velocity of the pulse wave. Compared to PTT and PWV, the PWA approach has a significant practical advantage as it uses a single PPG sensor to produce BP estimation. PWA techniques typically extract features from the PPG signal in the time and frequency domain and then use these features along with BP labels to train a regression model. Examples of machine learning algorithms used for this aim are linear regression [8-10], random forests [11], support vector machine (SVM) [12] and artificial neural network (ANN) [13-16]. The main drawback of this approach is that handcrafted features extracted from PPG signals are usually noise sensitive and may provide unreliable values. To mitigate this effect, a preprocessing stage is performed. This stage removes unreliable values and allows BP to be estimated more efficiently [12, 15, 16]. However, even with extensive preprocessing, previous attempts to estimate BP using this approach have mostly shown unsatisfactory results or been tested on a small dataset, making them difficult to generalize.

Recently, deep ANN learning methods have been successfully employed to deal with varied medical problems and in particular BP estimation from PPG [17-19]. A big advantage of this approach is that it does not require handcrafted features. Spectro-temporal PPG spectrograms are used as an input to ANN that is trained to automatically compute relevant PPG features and to use them for estimating BP values. [17, 18] use deep recurrent neural networks (RNNs) to learn long-term and short-term characteristics of a patient's PPG data. The RNN in [17] was trained on a small dataset so is difficult to generalize. In [18], a larger dataset is used but additional calibration data is required for each patient (e.g., age, gender, BMI and height). In [19], a relatively large dataset is used to train a CNN that estimates BP values using PPG signals only. The authors report MAD in the diastolic and systolic BP of 9.43 mmHg and 6.88 mmHg respectively. Such as error is unsatisfactory for most medical applications.

In this paper, we introduce two techniques for estimating diastolic and systolic BP values from PPG signals using CNNs. Section 2 presents the dataset we used and preprocessing performed on it to ensure reliable values. In Section 3 we describe the first technique that requires no calibration and in Section 4 we describe the second technique that achieves greater accuracy using Siamese network architecture that allows per-patient calibration. Finally, in Section 5, we discuss our results and conclude.

2. DATASET AND PREPROCESSING

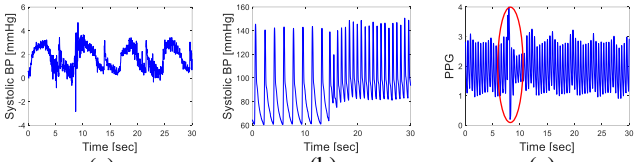
We used the MIMIC-II database [20] containing deidentified health-related data associated with thousands of patients who stayed within the critical care units of the Beth Israel Deaconess Medical Center between 2001 and 2008. This database contains a variety of vital medical parameters. We used two parameters for our dataset: PPG and correspondent arterial BP signals recorded at 125 Hz. We divided the PPG and BP signals to 30-second windows, as such a period represents stable BP values. This division yielded about 2.7 million such windows from 1,459 patients, representing more than 2.5 years of recordings. Data length varies from patient to patient, depending on their stay time in the critical unit. Since data from the MIMIC-II database are from patients in critical condition, they contain highly varied BP measurements as patients undergo treatment and receive drugs. The signals come from commercial devices and have been amplified and highly filtered.

Signals in this dataset often contain significant artifacts due to sensor fall off, sensor losing good contacts, etc. Other artifacts, related to respiratory activity, motion artifacts and system noise, are also common. The main artifacts we have identified are:

- Physiologically improbable BP values – Systolic BP values not in the range [75 165] or diastolic BP values not in the range [40 85]. Some signals even had negative values (Fig. 1(a)).
- Fluctuations in BP within a 30-second window – As systolic and diastolic BP values are determined by the mean of the maximum and minimum peaks of the BP signal respectively, signals with unstable peaks may have unreliable BP values (Fig. 1(b)).
- Noisy PPG and BP signals – Significant noise appears in many BP and PPG signals, making them unreliable. Note that even a minor noise (Fig. 1(c)) may corrupt the signal and significantly affect the accuracy of BP estimation.

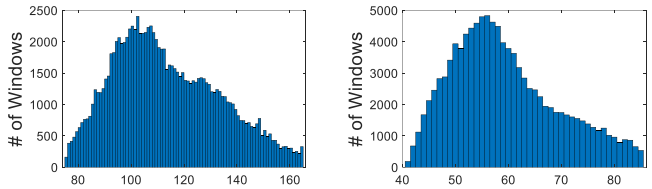
To obtain a reliable dataset, we must extensively preprocess the PPG and BP signals to remove windows with such artifacts. We do so by following these steps:

1. Removing unreliable windows – PPG and BP signals are periodic and are therefore typically characterized by high autocorrelation values. When corrupted by artifacts, these signals get lower autocorrelation values. To remove unreliable windows, we take similar steps for both PPG and BP. First, we remove the DC component. We then compute the autocorrelation for the window under consideration and normalize the result. Finally, we threshold the area under the squared magnitude of the computed autocorrelation so that low value windows,



(a) (b) (c)

Fig. 1. Example of BP and PPG artifacts in the MIMIC-II dataset. (a) BP signal with negative values. (b) Signal with unstable BP values. (c) PPG signal with minor noise (marked by red oval line).



(a) (b)

Fig. 2. Histogram of (a) systolic and (b) diastolic BP values in our clean MIMIC-II dataset.

corresponding to low periodicity, either of the PPG signal or the BP signal, are removed.

2. Removing unreliable patients - After removing unreliable windows as described above, many patients remain with only a few reliable windows. This may indicate unreliable signals from these patients, so we remove patients with less than 100 windows or that over 95% of their data removed in the previous step.
3. Outlier removal – Highly abnormal BP values are rare and therefore a reliable deep neural network estimator cannot be trained for these values. Furthermore, their accurate estimation is usually less clinically important. Therefore, for each patient, we remove all windows with BP values that vary over ±40 mmHg from the patient's first window that passed the previous two steps.

After these steps, we remain with a dataset containing 106,074 30-second windows belonging to 304 different patients. That is less than 5% of all data but of adequate quality and big enough for training CNNs. Even in these adequate quality windows, minor artifacts exist that can disrupt the extraction of systolic and diastolic BP values based on the maximum and minimum peaks, respectively. We address these artifacts by excluding peaks that are not close enough to the median value of the maximum or minimum peaks in each window. Fig. 2 depicts histograms of the systolic and diastolic BP values in our clean dataset. We randomly divided the clean dataset into 60% training set, 20% validation set and 20% test set. Unlike many other works in this field, great attention was paid to separating patients (and not windows) across the three sets. Namely, each set contains samples from different patients to prevent overfitting and to allow generalization of the trained models.

3.CALIBRATION-FREE BP ESTIMATION

We use a CNN to extract spectro-temporal features from PPG spectrograms obtained from each of the 30-second windows. Spectrograms are computed using the short-time Fourier transform (STFT) of sub-windows of length 6 seconds with a Hamming window and with 95% of overlap between adjacent sub-windows. We use a CNN architecture inspired by AlexNet [21], consisting of 5 convolutional layers and 3 fully connected layers. The first two convolutional layers and the fifth convolutional layer are followed by max pooling layers, while the third and fourth convolutional layers are connected directly. The output of these feature extraction layers goes into a series of two fully connected layers. Finally, unlike the original AlexNet architecture, the second fully connected layer feeds into a linear regression layer. ReLU activation function is applied after all convolution and fully connected layers. For regularization, a dropout layer precedes the first two fully connected layers, and we added a batch normalization layer following each convolutional layer. Another important parameter that we adjusted to our task was the loss function, from cross-entropy, as is common in classification networks, to L_1-loss, which is much more suitable for solving regression problems and that aim at minimizing the MAD used as a measure of BP estimation accuracy. The network was trained with the Adam optimizer [22] and a batch size of 32.

The results of this trained regression model are given in Table 1. Fig. 3(a) shows the confusion matrix of systolic BP estimation and Fig. 3(b) shows the confusion matrix of diastolic BP estimation. As expected, both confusion matrices are roughly diagonal. BP values with fewer data samples are farther from the diagonal which corresponds to higher MAD values.

4. BP ESTIMATION WITH SIAMESE NETWORK

Several previous works claim that per-patient calibration is crucial for accurate BP estimation from PPG as the BP signal depends on the specific properties of each patient's cardiovascular system. These works perform calibration using additional data such as patient age [9, 10, 18], gender, BMI and height [18]. We propose calibration using a single, first available 30-second window of PPG signal and its associated BP reading. The rationale behind this technique is that the physiological characteristics of each patient are embedded in his PPG spectrograms. To calibrate this way for each patient, we use a Siamese network architecture.

Siamese networks are neural networks that contain two identical subnetwork components, meaning that Siamese networks use the same architecture and parameters while working in tandem on two different input vectors to compute comparable output vectors and measure the distance between them. This architecture has been used in applications such as face recognition [23], signature verification [24], and matching queries with indexed documents [25].

Fig. 4 depicts our Siamese network architecture. The two identical CNN subnetworks have the same architecture

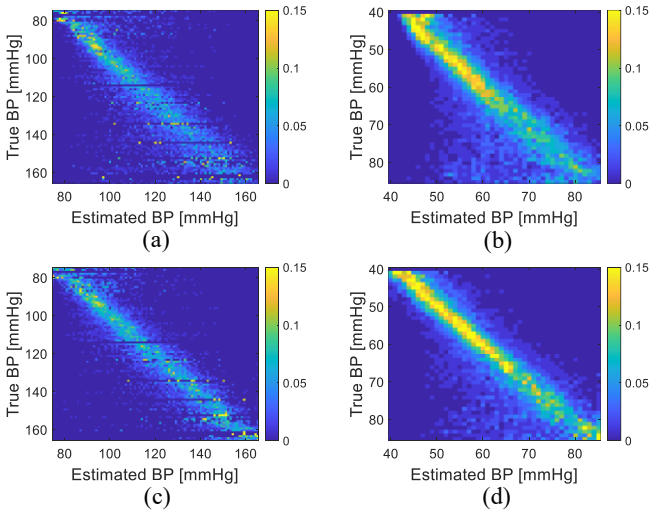


Fig. 3. Confusion matrices of the two proposed BP estimation techniques (a)(b) with no calibration and (c)(d) with Siamese network architecture that allows per-patient calibration for (a)(c) systolic and (b)(d) dialstolic BP estimation.

inspired by AlexNet described in Section 3, one receiving as input the current PPG spectrogram and the other receiving as input the PPG spectrogram of the first available window of the same patient. The only difference between this CNN architecture and the one used in the calibration-free technique described in Section 3 is that calibration-free CNN uses its final fully connected layer to deal with BP estimation as a regression problem, whereas here the result of each CNN is a feature vector. The two feature vectors are subtracted to yield a calibrated feature vector. If both input PPG spectrograms have close BP values, then their feature vectors are also expected to have similar values, whereas if the two input PPG spectrograms have distant BP values, then their feature vectors values are also expected to have very different values. Finally, the calibrated feature vector is fed to a ReLU activation layer and a linear regression layer to estimate the difference in BP values from the BP value of the calibration window. The hyper-parameters we use in this architecture are the same as listed in Section 3. The proposed architecture has two major differences compared to a classical Siamese network. First, it is a regression network. Second, while most Siamese networks use a metric with only positive values to estimate the distance between the two feature vectors corresponding to the two inputs, such as the Euclidian distance, we subtract the two vectors from each other without taking the absolute value to allow negative values as we also need to know the "direction" of the distance indicating higher or lower BP values compared to the BP value of the calibration window.

The results of the Siamese regression network described are given in Table 1. Fig. 3(c) shows the confusion matrix of systolic BP estimation and Fig. 3(d) shows the confusion matrix of diastolic BP estimation. Since estimation is more

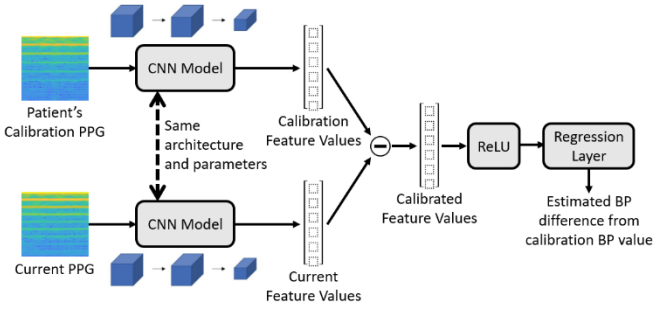


Fig. 4. Siamese network architecture for BP estimation.

accurate than that of the calibration-free technique, the confusion matrices have more values close to the diagonal.

5. DISCUSSION AND CONCLUSIONS

In this paper, we propose two techniques for estimating BP only from PPG using CNNs by representing short windows of PPG signals as spectrograms. One technique does not require calibration and the other estimates the change in BP values with respect to the patient's PPG and ground truth BP values at the time of calibration. To this aim, it uses a Siamese network architecture. The results of the two techniques on the clean MIMIC-II database are summarized in Table 1. To obtain these results, we had to preprocess the data extensively to remove unreliable values, as described in Section 2. The higher accuracy diastolic BP estimation versus systolic BP estimation is explained by the lower variance of the diastolic BP, as seen in Fig. 2.

Table 1. Accuracy of the two proposed techniques for BP estimation from PPG on our test set from the clean MIMIC-II database. Accuracy is measured by the mean absolute difference (MAD) and the standard deviation (STD).

	Calibration-free		**Siamese network**	
	MAD [mmHg]	STD [mmHg]	MAD [mmHg]	STD [mmHg]
Systolic BP	7.34	8.65	5.95	6.69
Diastolic BP	3.91	4.48	3.41	3.97

The AAMI recommends that the MAD of non-invasive BP measurements should not exceed 5 mmHg and the STD should not exceed 8 mmHg from a reference method. With calibration, we meet this requirement for diastolic BP and almost meet it for systolic BP. Even without calibration, the accuracy achieved is very promising. Since home BP monitors also do not meet these requirements in many cases, we can conclude that the results we obtained are sufficient for many practical medical applications.

6. ACKNOWLEDGMENT

The authors would like to thank the staff of Signal and Image Processing Lab. (SIPL) at the Technion and especially Prof. David Malah and Nimrod Peleg, for their comments and continuous assistance and support.

7. REFERENCES

[1] "ANSI/AAMI/ISO 81060-2: 2013 Non-Invasive Sphygmomanometers-Part 2: Clinical Investigation of Automated Measurement Type," Association for the Advancement of Medical Instrumentation, Arlington, VA, 2013.

[2] C. E. Evans, R. B. Haynes, C. H. Goldsmith, and S. A. Hewson, "Home blood pressure-measuring devices: a comparative study of accuracy," *Journal of Hypertension,* vol. 7, no. 2, pp. 133-142, 1989.

[3] D. Pendick, "Some home blood pressure monitors aren't accurate," 2014. Available: https://www.health.harvard.edu/blog/home-blood-pressure-monitors-arent-accurate-201410297494

[4] J. Allen, "Photoplethysmography and its application in clinical physiological measurement," *Physiological Measurement,* vol. 28, no. 3, p. R1, 2007.

[5] 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?," *Journal of Clinical Medicine,* vol. 7, no. 10, p. 316, 2018.

[6] M. Elgendi, "Optimal signal quality index for photoplethysmogram signals," *Bioengineering,* vol. 3, no. 4, p. 21, 2016.

[7] H. Tjahjadi and K. Ramli, "Review of photoplethysmography based non-invasive continuous blood pressure methods," in *15th International Conference on Quality in Research (QiR): International Symposium on Electrical and Computer Engineering*, 2017, pp. 173-178: IEEE.

[8] X. Teng and Y. 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*, 2003, vol. 4, pp. 3153-3156.

[9] R. Samria, R. Jain, A. Jha, S. Saini, and S. R. Chowdhury, "Noninvasive cuffless estimation of blood pressure using Photoplethysmography without electrocardiograph measurement," in *IEEE Region 10 Symposium*, 2014, pp. 254-257.

[10] S. Suzuki and K. Oguri, "Cuffless and non-invasive systolic blood pressure estimation for aged class by using a photoplethysmograph," in *30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2008, pp. 1327-1330.

[11] E. Monte-Moreno, "Non-invasive estimate of blood glucose and blood pressure from a photoplethysmograph by means of machine learning techniques," *Artificial Intelligence in Medicine,* vol. 53, no. 2, pp. 127-138, 2011.

[12] 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*, 2017, pp. 128-132.

[13] Y. Kurylyak, F. Lamonaca, and D. Grimaldi, "A Neural Network-based method for continuous blood pressure estimation from a PPG signal," in *IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, 2013, pp. 280-283.

[14] J. C. Ruiz-Rodríguez *et al.*, "Innovative continuous non-invasive cuffless blood pressure monitoring based on photoplethysmography technology," *Intensive Care Medicine,* vol. 39, no. 9, pp. 1618-1625, 2013.

[15] A. Gaurav, M. Maheedhar, V. N. Tiwari, and R. Narayanan, "Cuff-less PPG based continuous blood pressure monitoring - A smartphone based approach," in *38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2016, pp. 607-610.

[16] X. Xing and M. Sun, "Optical blood pressure estimation with photoplethysmography and FFT-based neural networks," *Biomedical Optics Express,* vol. 7, no. 8, pp. 3007-3020, 2016.

[17] C. Sideris, H. Kalantarian, E. Nemati, and M. Sarrafzadeh, "Building continuous arterial blood pressure prediction models using recurrent networks," in *IEEE International Conference on Smart Computing (SMARTCOMP)*, 2016, pp. 1-5.

[18] X. Li, S. Wu, and L. Wang, "Blood pressure prediction via recurrent models with contextual layer," in *Proceedings of the 26th International Conference on World Wide Web*, 2017, pp. 685-693.

[19] 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, 2019.

[20] M. Saeed *et al.*, "Multiparameter Intelligent Monitoring in Intensive Care II (MIMIC-II): a public-access intensive care unit database," *Critical Care Medicine,* vol. 39, no. 5, p. 952, 2011.

[21] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems (NIPS)*, 2012, pp. 1097-1105.

[22] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint: 1412.6980,* 2014.

[23] S. Chopra, R. Hadsell, and Y. LeCun, "Learning a similarity metric discriminatively, with application to face verification," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2005, pp. 539-546.

[24] J. Bromley, I. Guyon, Y. LeCun, E. Säckinger, and R. Shah, "Signature verification using a" siamese" time delay neural network," in *Advances in Neural Information Processing Systems (NIPS)*, 1994, pp. 737-744.

[25] P. Neculoiu, M. Versteegh, and M. Rotaru, "Learning text similarity with siamese recurrent networks," in *Proceedings of the 1st Workshop on Representation Learning for NLP*, 2016, pp. 148-157.