Towards Remote Blood Pressure Estimation Using RGB Cameras

Matthieu Scherpf, Hannes Ernst, Hagen Malberg, Martin Schmidt

Institute of Biomedical Engineering, TU Dresden, Dresden, Germany

Abstract

Arterial blood pressure (ABP) is one of the most important vital signs for the assessment of the human cardiovascular system. ABP is mainly characterized by the systolic (SBP) and the diastolic blood pressure (DBP). Early detection of elevated ABP is of utmost importance to prevent manifestation of severe diseases such as stroke. In order to achieve regular self-monitoring of ABP, measurement devices should be as convenient as possible. This work investigates a novel approach based on RGB cameras to estimate ABP remotely. We applied imaging photoplethysmography (iPPG) for the extraction of the blood volume pulse (BVP) from 56 healthy subjects with induced elevated SBP. The BVP was then used as the input for a customized neural network (Vnet) operating in time and frequency domain. SBP and DBP were estimated based on 20-second segments. To evaluate our approach, nested cross validation was applied. We then calculated the Pearson correlation coefficient (r) to quantify model agreement for SBP and DBP estimation. For generalized ABP estimation, Vnet achieved r of 0.30 and 0.19 for SBP and DBP, respectively. In contrast, for personalized ABP estimation, r reached 0.71 and 0.51 for SBP and DBP. Our work demonstrates the potential of ABP estimation based on iPPG and delivers important insights into the relevance of optimal model selection from cross validation.

1. Motivation

Arterial blood pressure (ABP) is one of the most important vital signs for the assessment and the early detection of deterioration of the cardiovascular system [1]. ABP is mainly characterized by the systolic blood pressure (SBP) and diastolic blood pressure (DBP) representing the maximum and minimum pressure in the arterial system, respectively. A beginning deterioration is most often diagnosed by the continuous increase of SBP leading to hypertension. SBP and DBP are commonly measured in an auscultatory or oscillometric manner using an inflatable cuff which requires either medical staff or a specific medical device. This kind of measurement is cumbersome and prone to errors due to artifacts, e.g. resulting from the white coat syndrome or incorrect cuff placement.

Several methods for remote estimation of SBP and DBP based on imaging photoplethysmography (iPPG) have been previously presented [2]. However, due to unclear dataset splitting, comparability is impeded and the quality of validation is often insufficient. Within this work, we propose a novel approach for remote ABP measurement based on RGB cameras using iPPG. A custom neural network (Vnet) is applied for SBP and DBP estimation, which combines features from time and frequency domain. We evaluated our approach on 56 healthy subjects from the "Dresden Multimodal Biosignal Dataset for the Mannheim Multicomponent Stress Test" (DMBD-MMST) [3] with overall 2300 minutes of uncompressed video and synchronized continuous non-invasive ABP recordings. The participants underwent a mental stress test followed by a relaxation phase, which first triggered ABP to increase and subsequently decrease. By applying Vnet to the DMBD-MMST we (1) aimed to examine the impact of different data splitting and model selection methods for SBP and DBP estimation and (2) investigated the potential for automated detection of elevated SBP.

2. Methodology

We extracted remote BVP based on iPPG from 20-second RGB video recordings from the facial skin region. The BVP was then preprocessed and used as input for a custom convolutional neural network, further denoted as Vnet (due to its shape), which estimated SBP and DBP. The processing pipeline for SBP and DBP estimation based on iPPG and Vnet's architecture are shown in Fig. 1.

2.1. Data Description

We used the DMBD-MMST dataset comprising 65 healthy subjects who underwent a mental stress test [3]. One experiment lasted around 2 hours and consists of 6 consecutive uncompressed 5 to 10-minute RGB video recordings acquired with an industrial camera at a resolution of 640 pixels high and 320 pixels wide at 100 frames per second and 12-bit color depth. Synchronously, the reference SBP and DBP for each cardiac cycle was acquired using a Task Force Monitor 3040i from CNSys-

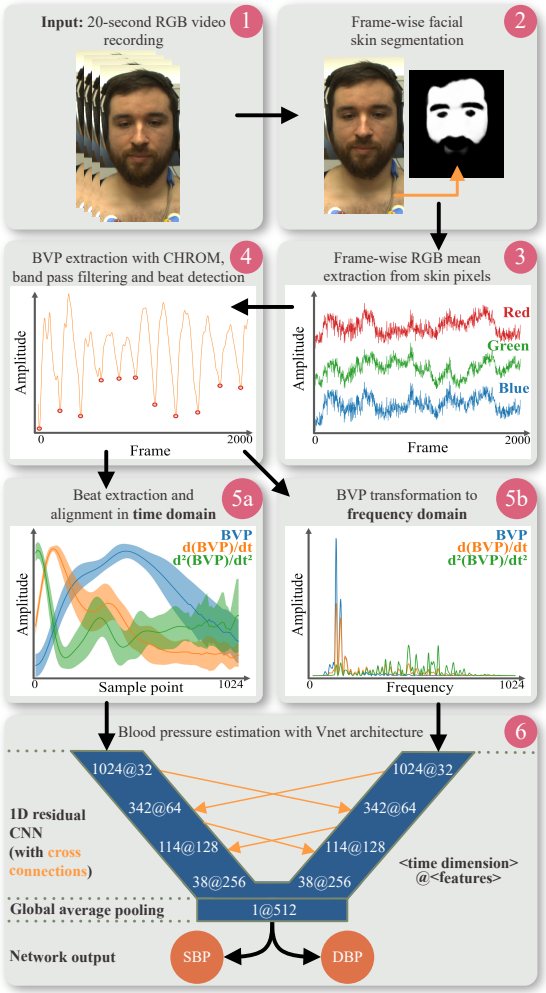

Figure 1. Processing pipeline to estimate systolic (SBP) and diastolic blood pressure (DBP) from a 20-second RGB video recording using Vnet, a customized convolutional neural network (CNN). In 5b, the solid lines and the area around it indicate the mean and standard deviation for the aligned beats.

tems Medizintechnik GmbH (Graz, Austria). After artifact removal and elimination of erroneous recordings, the dataset comprised of recordings from 56 subjects. A complete 5 to 10-minute recording was deemed to be erroneous and excluded in case of synchronization inconsistencies. After splitting the recordings into segments of 20 seconds length, a segment was excluded if the signal-to-noise-ratio (SNR) of the BVP, computed as in [4], was lower than 3 dB, primarily due to significant head movement, or if arithmetic mean of SBP or DBP were lower than 90 mmHg or 60 mmHg, respectively. These values are defined as hypotension [5] and hence were classified as abnormal. In contrast, we did not exclude elevated SBP and DBP measurements because their occurrence was expected within the design of the experiment and were investigated in this study. For the filtered dataset, the mean and standard deviation of SBP and DBP was 114 ± 12 mmHg and 74 ± 8 mmHg, respectively.

2.2. BVP Extraction

For BVP extraction, we first segmented the facial skin pixels for each frame (Fig. 1.2) using a similar approach as in [4], but replaced the DeepLabV3+ network with a Unet architecture, which has improved efficiency and accuracy. Next, the arithmetic mean of all skin pixels for each color channel and each frame was calculated (Fig. 1.3). Subsequently, we computed the BVP with CHROM which we selected because of its comprehensively validated high performance regarding BVP extraction (Fig. 1.4) [6]. The BVP was further filtered using a zero-phase 4th order Butterworth filter. Because the average heart rate was 71.3 ± 12.7 beats per minute, the pass band ranged from 0.67 Hz to 4 Hz as we expected most blood pressure related information in the spectral range around 1 to 3 times the heart rate.

2.3. BP Estimation

As shown in Fig. 1, from each segment time and frequency domain inputs were computed. We extracted the beats for each segment, applied min-max normalization for each beat, aligned them and computed the mean and standard deviation (Fig. 1.5a). This was done for the BVP, its first and second derivative. Additionally, we computed the spectral components up to 4 Hz of each segment using Fast Fourier Transformation (Fig. 1.5b). Reference SBP and DBP were averaged over all cardiac cycles in one segment. To process the computed inputs, extract and combine meaningful features for blood pressure estimation, we implemented Vnet (Fig. 1.6), a two-branched architecture. The left branch processes the time domain, i.e. the aligned beats, and the right branch processes the spectral components. Vnet was setup with residual connections and in addition, we implemented cross connections to enhance the information exchange and enable a flexible feature combination between the 4 subsequent abstraction layers of the left and right branch. We trained Vnet with AdamW optimizer, a dynamic learning rate starting at $3 \cdot 10^{-4}$ and decreasing exponentially to $3 \cdot 10^{-5}$ and a batch size of 32. To mitigate overfitting, we added dropout with a rate of 50 %. We set up the convolutional layers with a kernel size of 5 and a stride of 3. Vnet was trained to minimize mean squared error (MSE) with equal weights for SBP and DBP and trained for 160 epochs.

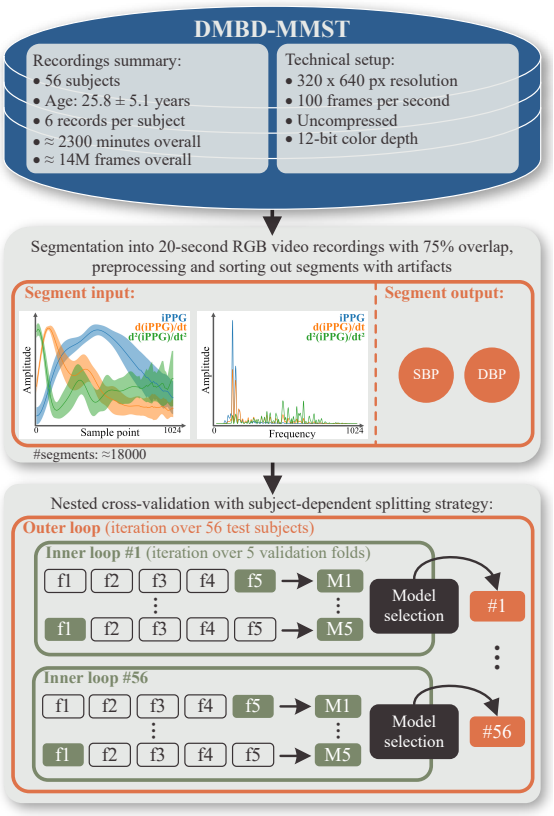

Figure 2. Dataset description, dataset segmentation and nested cross validation procedure. f1 to f5 and M1 to M5 indicate the folds and the trained models. The validation folds are filled in green.

2.4. Data Splitting and Model Selection

Due to subject specific physiological parameters and characteristics, e.g. the influence of blood vessel stiffness [1], ABP estimation from iPPG-based BVP is a complex task requiring a considerable amount of data to achieve meaningful generalization using machine learning approaches. With 56 subjects our dataset is considered small. To maximize the training data and to enable valid model selection and robust evaluation, we applied a nested cross validation illustrated in Fig. 2. For the outer cross validation, we applied 56 times leave one subject out. For the inner cross validation, i.e. the remaining 55 subjects, a 5-fold cross validation was used with 11 subjects per fold. Therefore, each model was trained on 44 subjects and validated on 11 subjects. We trained 5 models per subject and 280 models overall. For each model, we used the state from the epoch with lowest MSE on the corresponding validation fold. A significant step for the final estimation of the segment-based SBP and DBP for each of the test subjects is the model selection. We applied 3 different strategies for the final SBP and DBP estimation based on the 5 models for each of the test subjects. The first strategy used 5 models and applied ensemble averaging over each of the 5 estimations for each segment. The second strategy used only the best of the 5 models based on lowest average MSE on the validation fold. The third strategy aimed at determining the highest possible overall performance by selecting the best of the 5 models based on the lowest average MSE on the test subject representing a personalized selection approach.

Table 1. Results for the 3 model selection strategies for systolic (SBP) and diastolic blood pressure (DBP) estimation based on Pearson correlation coefficient (r) and root mean squared error (RMSE). F1 score was used for performance evaluation regarding elevated SBP detection.

#	Model selection strategy	Metric	SBP	DBP
1	Ensemble averging	r	0.08	0.12
2	Best of 5 on val. folds		0.30	0.19
3	Personalized		**0.71**	**0.51**
1	Ensemble averging	RMSE (mmHg)	11.45	8.04
2	Best of 5 on val. folds		12.18	8.69
3	Personalized		**7.55**	**6.76**
1	Ensemble averging	F1 score	0.01	-
2	Best of 5 on val. folds		0.18	-
3	Personalized		**0.35**	-

2.5. Evaluation

We computed segment-based SBP and DBP for all test subjects with the corresponding model(s) based on the selection strategy. Pearson correlation coefficient (r) and root mean squared errors (RMSE) were used to evaluate the overall model performance. F1 score was used to quantify the detection performance of elevated SBP, which was labeled for segments with SBP higher than 130 mmHg.

3. Results and Discussion

The results are presented in Table 1. The ensemble averaging approach achieved the lowest performance with r around 0.1 and F1 score of 0.01. If the best of 5 models regarding MSE on the validation fold were selected, r reached 0.3 and 0.19 for SBP and DBP, respectively, and F1 score of about 0.18. The personalized model selection approach yielded the best results with r of 0.71 (shown in Fig. 3) and 0.51 for SBP and DBP, respectively, and F1 score of 0.35. Overall, SBP estimation achieved higher performance in r and RMSE than DBP estimation. RMSE was lowest for the personalized model selection approach.

We assume that the ensemble averaging approach achieves the lowest results because some models tend to exhibit very low performance on the corresponding test subjects and therefore their estimates significantly degrade the averaged model estimates. We investigated this by using the personalized model selection approach selecting 3

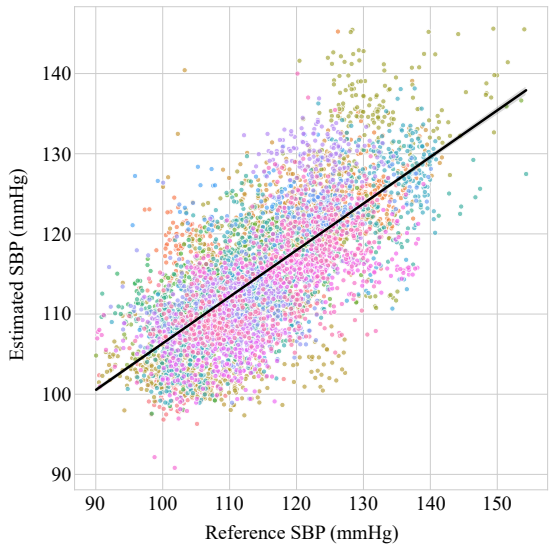

Figure 3. Estimated and reference systolic blood pressure (SBP) of the personalized model selection approach for all 56 test subjects indicated by different colors. The black line represents the linear regression.

best out of 5 models and still achieved r of 0.60 and 0.43 for SBP and DBP, respectively. Therefore, selecting the best of 5 models based on the validation results is prone to suboptimal model selection due to the small validation fold of only 11 subjects. The personalized approach indicates the existence of a model selection strategy that significantly improves generalized ABP estimation based on the DMBD-MMST. Conventional methods typically require calibration to enable accurate remote ABP estimation. In contrast, the personalized approach presented here offers a new perspective by identifying the model that accurately represents a subject's characteristics. Due to the small fraction of elevated SBP segments, the models tend to underestimate SBP and therefore F1 scores are generally low, even for the personalized model. Fig. 3 shows the estimated and reference SBP values and reveals this bias in the estimations with a regression line that is less steep than the 45° line.

4. Conclusion

Our results indicate that ABP estimation based on iPPG is possible to a certain performance level depending on the underlying model selection strategy. We identified two key findings from our investigation: First, ABP is a complex physiological parameter which is why a larger dataset with increased variation in ABP range is necessary for improved generalization performance. Second, future research should further investigate possible heuristics for the model selection process as our results indicate opportunities for improvement. Furthermore, our personalized selection approach provides insights for additional ABP calibration methods based on iPPG.

We investigated remote detection of elevated SBP based on 20-second segments. The maximum F1 score of 0.35 for the personalized approach is insufficient for practical use. However, we anticipate significant improvements in the future, particularly with a more diverse dataset.

Since signal quality is of great importance, future research should also focus on more precise iPPG-based BVP extraction. Advanced approaches, such as the deep learning-based DeepPerfusion model [7], enable high-precision BVP extraction and should therefore be more suitable for blood pressure estimation.

Acknowledgments

This research was partially funded by the German Federal Ministry of Education and Research (BMBF) as part of the research program "Interactive technologies for health and quality of life" (project KOMVITAL, grand number 16SV9245). We acknowledge the computing time made availabel by the NHR Center of TU Dresden.

References

[1] Avolio AP, et al. Arterial blood pressure measurement and pulse wave analysis–their role in enhancing cardiovascular assessment. Physiological Measurement 2010;31(1):R1–47. ISSN 0967-3334.

[2] Schrumpf F, et al. Assessment of non-invasive blood pressure prediction from ppg and rppg signals using deep learning. Sensors 2021;21(18). ISSN 1424-8220.

[3] Ernst H, et al. Assessment of the human response to acute mental stress-an overview and a multimodal study. PloS one 2023;18(11).

[4] Scherpf M, et al. Skin segmentation for imaging photoplethysmography using a specialized deep learning approach. In 2021 Computing in Cardiology (CinC). IEEE. ISBN 978-1-6654-7916-5, 2021; 1–4.

[5] Sharma S, et al. StatPearls: Hypotension. Treasure Island (FL), 2024.

[6] de Haan G, et al. Robust pulse rate from chrominance-based rppg. IEEE transactions on bio medical engineering 2013; 60(10):2878–2886.

[7] Scherpf M, Ernst H, Malberg H, Schmidt M. Deepperfusion: A comprehensible two-branched deep learning architecture for high-precision blood volume pulse extraction based on imaging photoplethysmography. TechRxiv, 2024. Preprint.

Address for correspondence:

Matthieu Scherpf
Institute of Biomedical Engineering, TU Dresden
Fetscherstraße 29, 01307 Dresden
Matthieu.Scherpf@tu-dresden.de