Blood Pressure Estimation from Photoplethysmography Using Hybrid Scattering–LSTM Networks

Osama A. Omer (omer.osama@aswu.edu.eg)
Aswan University

Mostafa Salah
Sohag University

Ammar M. Hassan
Arab Academy for Science, Technology and Maritime Transport

Norihiro Sugita
Tohoku University

Yoshifumi Saijo
Tohoku University

Research Article

Keywords: Photoplethysmography (PPG), continuous blood pressure monitoring, Arterial Blood Pressure (ABP), LSTM, Deep Learning, Wavelet Scattering Transform, Regression

Posted Date: February 6th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2538573/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Abstract

One of the most significant indicators of heart and cardiovascular health is blood pressure (BP). Blood pressure (BP) has gained great attention in the last decade. Uncontrolled high blood pressure increases the risk of serious health problems, including heart attack and stroke. Recently, machine/deep learning is leveraged for learning the BP from Photoplethysmography (PPG) signals. Hence, continuous BP monitoring can be introduced based on simple wearable contact sensors or even remotely sensed from a proper camera away from the clinical setup. However, the available training dataset imposes many limitations besides the other difficulties related to the PPG time series as high-dimensional data. This work presents beat-by-beat continuous PPG-based BP monitoring while accounting for the aforementioned limitations. For a better exploration of beats’ features, we propose to use wavelet scattering transform as a better descriptive domain to cope with the limitation of the training dataset and to help the deep learning network accurately learn the relationship between the morphological shapes of PPG beats and the BP. Long Short-Term Memory (LSTM) network is utilized to demonstrate the superiority of the wavelet scattering transform over others domains. The learning scenarios are carried out on a beat basis where the input corresponding PPG beat is used for predicting BP in two scenarios; 1) Beat-by-beat arterial blood pressure (ABP) estimation, and 2) Beat-by-beat estimation of the systolic and diastolic blood pressure values. Different transformations are used to extract the features of the PPG beats in different domains including time, DCT, DWT, and wavelet scattering Domains. The simulation results show that using the wavelet scattering domain outperforms the other domains in the sense of root mean square error (RMSE) and mean absolute error (MAE) for both of the suggested two scenarios.

1. Introduction

As the control of blood pressure (BP) is so important, monitoring BP regularly has become a critical concern. According to the World Health Organization (WHO) [1], 17.9 million people die each year from cardiovascular illnesses. The American Society of Anesthesiologists (ASA) has recommended basic anesthetic monitoring criteria [2]. According to ASA, blood pressure should be taken every 5 minutes to ensure that the patient’s circulatory system is functioning properly. Continuous arterial blood pressure (ABP) represents the gold standard of BP monitoring, however, it is difficult to measure in routine clinical practice because it is performed via arterial cannulation by a trained operator. Auscultatory, oscillometry, tonometry, and volume clamping methods are commonly used in traditional BP measurement. Traditional blood pressure monitoring techniques have two main drawbacks: first, they require special instruments that may not be available to everyone. The second one is that these methods are not always suitable for babies and elder people. As a result, using existing technologies is challenging since they are unsafe to use at a high rate. Hence, continuous BP monitoring cannot be carried out through traditional non-invasive approaches. Usually, BP is represented through three essential readings including systolic, diastolic, and mean pressure, which are defined as follows. The pressure imposed on the walls of blood arteries by blood at the conclusion of systolic contraction of the ventricles is known as systolic blood pressure (SBP). The pressure imposed on the blood vessel walls at the conclusion of diastole, or
relaxation, is known as diastolic blood pressure (DBP). The pressure that determines the average rate of blood flow through systemic arteries is known as mean pressure.

Recently, Photoplethysmography (PPG) signal is exploited for providing proper continuous BP monitoring [3–5] through simple wearable devices [6]. PPG signal represents changes in blood volume in the skin blood vessels by tracking fine variations in the absorbed/reflecting infrared light intensity due to blood pulsation. Actually, these changes not only explain blood volume but also provide information on blood flow and red blood cell orientation. Thanks to the publically available MIMIC II data set (Multi-parameter Intelligent Monitoring in Intensive Care) [7] that provides joint PPG-ABP data along electrocardiogram (ECG) signal. There are many research efforts devoted to inferring the relationship between PPG and BP based on these available datasets through machine learning [8–13] and deep-learning techniques [14–19]. Some examples of PPG-based BP estimation can be reviewed shortly as follows.

The authors in [13] extracted features from contact PPG to learn a neural network to determine the BP. This system has experimented on patients. In [14], deep learning networks have been evaluated for BP estimation based on both contact PPG and camera-based remote PPG (rPPG) signals. A spectro-temporal deep neural network is employed to estimate BP in [15]. The authors used the MIMIC III database to train the network after many prepossessing operations. The authors used the MIMIC II database [7] to train the network after many prepossessing operations. The experimental results showed that the mean absolute errors were 9.43 and 6.88 for SBP and DBP, respectively. Rather than a single PPG sensor approach, two sensors may be employed to estimate BP as in [17, 20–24]. In [23], the authors suggested using the pulse transit time (PTT) and pulse arrival time (PAT) with the PPG signal that is collected using the contact method to find the blood pressure. In [21], the authors confirmed the relationship between pulse wave velocity (PWV) and SBP by experiments applied to 63 volunteers. In [17], the authors have employed a two-stage U-net network for predicting continuous arterial blood pressure (ABP) using ECG and PPG signals together. ABP has been eventually used to calculate systolic blood pressure (SBP) and diastolic blood pressure (DBP) by finding the maximum and minimum values of the ABP signal, respectively. Jointly using two kinds of signals, ECG and PPG signals, in this method is not recommended because it needs signals from two different sensors. Also, the estimated signals in this scheme need more refinements to be clearer. In [25], a scheme to convert the PPG signal to the ABP signal has been introduced using the federated learning approach. The simulation results explained that the mean absolute error was improved to reach 2.54 mmHg. In contrast, the standard deviation was 23.7 mmHg for the mean arterial blood pressure.

**Challenges and Motivations**

- Data quality: Clean data is mainly required for fully exploiting deep-learning power. However, any practical dataset needs some cleaning for excluding deformed signals, hence, mitigating disturbing the training operations. Strict data cleaning reduces the amount of available data essentially to the extent that becomes insufficient for the network for inferring the underlying phenomena.
• PPG nature as time series: The nature of the PPG signal can be described as a time series or high dimensional data. The different signals exhibit general deformation and translation. Moreover, based on our observations, there are many PPG classes or shapes related to the same BP range. So, learning these complicated relationships requires effective feature simplification before the deep learning stage.

Contributions

For a better exploration of beats’ features, we propose to merge deep learning with efficient feature extraction means to cope with the limitation of the training dataset and to help the deep learning network accurately learn the relationship between the morphological shapes of PPG beats and the BP. Hence, deep learning is applied to a better representative domain rather than the time domain. So, we can summarize our contributions as follows:

• Wavelet scattering transform (WST) is applied for providing an efficient feature extraction prior deep earning stage. It adds shift invariance property along with some resistance to local deformation.
• The performance of wavelet scattering transform is compared against different feature extraction domains including time domain, discrete wavelet transform (DWT), and discrete cosine transform (DCT).
• The learning scenarios are carried out on a beat basis where the input PPG beat is used for predicting:

  a. Corresponding complete ABP beat.
  b. Just, systolic and diastolic BP values.

2. Materials Used

This section describes the used dataset employed in the proposed deep-learning scenarios. Also, the preprocessing operations are described as well.

2.1 Data Set

The combined PPG-ABP data, which is needed as input to the learning models, is provided by Physionet’s MIMIC II data collection (Multi-parameter Intelligent Monitoring in Intensive Care) [26]. In [17], a better-assembled version of the data is provided publically. An inspection of that data set, however, indicates a high number of erroneous PPG and ABP signals. That dataset will be used to produce a simultaneously cleaned PPG-ABP dataset [27] that will be fed into two deep learning-based BP estimation models. This original data set contains 12,000 records of varying lengths. In each record, ABP (invasive arterial blood pressure in millimeters of mercury), PPG (photoplethysmograph from fingertip), and ECG (electrocardiogram from channel II) data are taken at Fs = 125 samples per second. However, we're just
interested in the PPG beats and the ABP labels/beats that go with them. For efficient processing and filtering, records are separated into 1024 sample pieces. So far, 30,660 signals have been introduced.

2.2 Preprocessing

PPG signals can only be pre-processed using any boosting method (for example, band-pass filtering in the [0.5-8] Hz frequency range) as long as their morphological form is not altered. On the other hand, the ABP signal cannot be manipulated since any attempt to increase its quality changes its magnitude, which shows the BP value. During the beat segmentation step, PPG beats will also be altered, but corresponding ABP beats will be kept alone. Without filtering, significantly distorted ABP signals or beats must be eliminated.

3. Methodology

In this work, continuous beat-by-beat BP monitoring is proposed based on wavelet scattering transform. We are motivated by wavelet scattering transform capabilities that meet the underlying challenges imposed by the usage of the available dataset. WST provides unsupervised feature extraction hence, the training is performed on the better representative feature domain rather than the noisy time domain. It is recommended to merge the ScatNet with conventional deep learning networks such as the Convolutional Neural Network (CNN) for enhancing its learning performance and stability [28–31]. However, for 1-D signals or time series modeling, recurrent Neural Network (RNNs) such as Long Short-Term Memory (LSTM) is preferred over CNNs. So, ScatNet is used in the first stage for simplifying signal features then followed by the LSTM network in the next stage for training on the simplified representations. Specifically, the ScatNet introduces trustworthy feature retrieval at various beat sizes along with invariance to shift/rotation and small deformation. Hence, it absorbs some of the encountered dataset artifacts.

We are introducing two BP prediction scenarios (shown in Fig. 1): 1) **Beat-by-beat cPPG/ABP mapping**: The training is performed for mapping individual PPG beats into corresponding ABP beats. Hence, SBP/DBP (systolic/diastolic BPs) can be found as the maximum/minimum values of the output ABP beat. 2) **Beat-by-beat cPPG-to-SBP/DBP mapping**: The training is performed for mapping individual contact PPG (cPPG) beats into corresponding SBP/DBP values directly. The two scenarios are performed on a beat basis for providing continuous beat-by-beat BP minoring. Hence, the PPG signal is segmented into beats and the time interval is recorded. Beat-by-beat selection is performed to select valid beats for training, testing, and validation. It worthies to note that the time indexes of all beats are normalized to be with a fixed length, however, the time interval of each beat is used as extra data in the training, testing, and validation processes in the case of time domain features. The corresponding ABP signals are segmented as well into beats, however, the ABP beats are normalized in time but not in amplitude. In the second scenario, the normalized PPG beats are used with the BP (SBP and DBP) for each beat for training, validation, and testing. The following processes are applied to estimate the BP in both scenarios. Moreover, the proposed system is evaluated with the remotely estimated PPG (rPPG) to estimate ABP and BP.
3.1 Preprocessing Stage

To remove the contaminating noise, PPG signals are passed through a bandpass filter in the cardiac frequencies [0.5 Hz – 8 Hz] where the fundamental frequency locates in this range a round fundamental frequency, and the first harmonics.

3.2 Signal Segmentation

The filtered signal is divided into beats in this stage so that each beat may be addressed separately. The identification of local minimum locations serves as the foundation for signal segmentation. The local minimum locations for the filtered signal are shown in Fig. 2. A crucial feature of the beats is the beat interval (BI), which is the period between each pair of successive minimums [32].

3.3 Beat Selection

Not all the beats are valid for training, testing, validation, and BP prediction. Therefore, three criteria are used for selecting valid beats. Figure 3 shows a plot for all the segmented beats of the filtered signal. In this figure, it can be shown that many beats are not valid, i.e. with irregular shapes. These beats may disturb the DL network in both training and testing processes.

To reject these beats, the following criteria are used [27]:

- **Beat interval (BI):** the standard range of the heart rate is [40 bpm – 180 bpm] which corresponds beat interval in the range [1.5–0.33] second. Therefore, we use only beats with beat intervals in the range $0.33 \leq BI \leq 1.5$.

- **Beat Skewness quality index (SQI):** Skewness is defined as the difference between the shape of the beat and the standard beat shape. The normal beats have positively skewed shapes. It is also can be called the right-skewed beat. A tail is referred to as the tapering of the curve differently from the data points on the other side If the given beat is shifted to the right and with its tail on the left side, it is a negatively skewed beat. It is also called a left-skewed distribution. The skewness value of any distribution showing a negative skew is always less than zero. SQI can be calculated as

$$ SQI = \frac{\sum_{i=1}^{N}(Y_i - Y^\sim)^3 / N}{S^3} \quad (1) $$

Where $Y^\sim$ is the mean, $S$ is the standard deviation, and $N$ is the number of beat's points. In this work, we restrict our beats with only positive SQI. Moreover, the high positivity of SQI leads to a long-tailed beat which is not a normal beat. In our work, we use valid beats with $0 \leq SQI \leq 1$.

- **Beat correlation quality index (CQI):** the correlation with a standard beat can help restrict beats to be within a valid range. However, the correlation should be not strict that is CQI should be more than 0.3
to ensure rejecting highly deviated beats. Figure 4 demonstrates some of the accepted beats according to these criteria.

3.4 Deep Learning Model

Wavelet scattering transform (WST), is introduced as an effective tool for feature extraction. It provides shift-invariant representation that is stable to rotation and local deformation [33]. In this way, it discards the uninformative variations in the signal. So, it has been employed for ECG signal classification [34]. It is regarded as a deep convolutional network (referred to as “ScatNet”) [35] that cascades wavelet transforms and modules nonlinearity. It addresses the main challenges related to CNN networks, represented in, 1) the need for large training data and computations, 2) the choice of many hyperparameters, and 3) the difficulty in interpreting the features. Rather than CNN networks, the filters employed in the convolution stage in ScatNet are found without training. Features can be explained and visualized efficiently [30]. The following subsections describe both WST and LSTM network. Figure 5 shows the proposed hybrid deep-learning model along different feature extractors in different domains.

3.5 Wavelet Scattering Transform

To compute the WST of a signal $x$, this signal is processed in three successive operations to generate wavelet scattering coefficients in each stage. These operations are convolution, nonlinearity, and averaging respectively. Figure 6 displays multiresolution/multilayer wavelet scattering transforms in which the scattering coefficients should be determined at each layer [35] as follows:

The zeroth-order scattering coefficients are calculated using basic input averaging as follows:

$$S_0 = x \star \phi,$$  \hspace{1cm} (2)

where $x$ is the input signal, $\phi$ is the scaling function, and $\star$ is the convolution operator.

The high-frequency bands are captured by convolving with the mother wavelet $\psi_{\lambda_1}$ at scale $\lambda_1$. So, the first-order scattering coefficients, $S_1$, are generated by averaging the modules of the lowest band at the first scale in the filter bank $\lambda_1$.

$$S_1 x(t, \lambda_1) = \mid x \star \psi_{\lambda_1} \mid \star \phi$$ \hspace{1cm} (3)

In the same way, the second wavelet transform is determined as follows:

$$S_2 x(t, \lambda_1, \lambda_2) = \mid \mid x \star \psi_{\lambda_1} \mid \star \psi_{\lambda_2} \mid \star \phi$$ \hspace{1cm} (4)

At each step, the signal at the lowest band incurs modules non-linearity and is averaged through the convolution by the father wavelet (low pass) $\phi$ filter as shown in fig.1. For the $m^{th}$ layer, the scattering coefficients, $S_m$, have been computed as follows.

$$S_m x(t, \lambda_1, ..., \lambda_m) = \mid \mid \mid x \star \psi_{\lambda_1} \mid \star ... \star \psi_{\lambda_m} \mid \star \phi$$ \hspace{1cm} (5)

3.6 LSTM network

The LSTM neural networks for integrated channel equalization and symbol detection are covered in this section. The proposed DL-LSTM-based BP and ABP estimate is trained offline with simulated data.
Input, output, and forget gates, as well as a memory cell, comprise the LSTM NN structure shown in Fig. 7. The LSTM NN properly stores long-term memory via the forget and input gates. The LSTM cell’s primary structure is depicted in Fig. 6, in [36]. The forget gate allows the LSTM NN to eliminate the unwanted data from the last process by using the Presently used input $x_t$ and the cell output $h_{t-1}$. On the basis of the preceding cell output $h_{t-1}$ and the Present cell’s input $x_t$, the input gate determines the data that will be utilized in conjunction with the preceding LSTM cell state $c_{t-1}$ to generate a new state of the cell $c_t$. LSTM may decide which data is discarded and which is maintained by using the forget and input gates. The LSTM network is a form of recurrent neural network (RNN) that is smart enough to learn long-term correlations between time series [36].

4. Experimental Results And Discussion

In this section, the results of BP from PPG are reported according to two learning scenarios: 1) Per-beat continuous PPG-to-ABP learning, and 2) Per-beat discrete PPG-to-SBP/DBP learning scheme. These scenarios are shown in Fig. 5, while the detailed network configurations are described in Tables 1 and 2. Based on these two learning scenarios, we are interested in studying the impact of the feature extraction stage on the overall performance. So, different feature extractors are used besides WST.

The cleaned dataset [27] is employed. The dataset was split into training, validation, and test sets on a beat basis to prevent contamination of the validation and test set by training data. We used 158094 beats for training, 17566 beats for validation, and 17566 beats for testing. We used the MAE metric to assess the performance of all methods. We determined the prediction errors for the full dataset.

The segmented PPG beats are normalized to be in the range [0–1] by using the following equation.

$$S_n = \frac{S - \min (S)}{\max (S) - \min (S)} \quad (6)$$

where $S_n$ is the normalized signal and $S$ is the un-normalized beat. The normalized beats are then normalized in time so as to be with a fixed length (120). The time interval is used as a feature besides the normalized beats. The corresponding ABP signals are also divided into beats, however, the ABP beats are only normalized in terms of time and not amplitude. The complete ABP beat represents the label for the first scenario, while only its maxima (SBP) and minima (DBP) points are utilized for labeling in the second scenario.

For the two tested scenarios, the input to the LSTM network is a beat (or corresponding feature domain representation) with a length of 120 in addition to BI (in the case of time domain features) as tabulated in Table 2. In the first scenario, ABP estimation, the output is a sequence with a size equal to $120 \times 1$ for all feature domains. On the other hand, in the second BP estimation scenario, the output is two values that are the SBP and the DBP, for all feature domains.
### Table 1
Network specifications

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of Beats</strong></td>
<td>158094</td>
</tr>
<tr>
<td><strong>Beat Length (time samples)</strong></td>
<td>120</td>
</tr>
<tr>
<td><strong>Input Feature Domain</strong></td>
<td>One of the following domains:</td>
</tr>
<tr>
<td></td>
<td>1- Time Domain.</td>
</tr>
<tr>
<td></td>
<td>2- DCT Domain.</td>
</tr>
<tr>
<td></td>
<td>3- DWT Domain.</td>
</tr>
<tr>
<td></td>
<td>4- WST Domain</td>
</tr>
<tr>
<td><strong>Number of Channels</strong></td>
<td>4×1 Layer array</td>
</tr>
<tr>
<td><strong>Layer specifications</strong></td>
<td>1- Sequence input with 120 dimensions</td>
</tr>
<tr>
<td></td>
<td>2- LSTM with 20 hidden units</td>
</tr>
<tr>
<td></td>
<td>3- Two fully connected layer</td>
</tr>
<tr>
<td></td>
<td>4- Regression Output mean-squared-error</td>
</tr>
<tr>
<td><strong>Learning Rate</strong></td>
<td>0.005</td>
</tr>
<tr>
<td><strong>Number of Iterations per Epoch</strong></td>
<td>191</td>
</tr>
<tr>
<td><strong>Optimization function</strong></td>
<td>L2-Norm</td>
</tr>
<tr>
<td><strong>Optimization method</strong></td>
<td>ADAM</td>
</tr>
</tbody>
</table>

### Table 2
Input and output data size in the two used scenarios for different domains.

<table>
<thead>
<tr>
<th>Per-beat Scenario</th>
<th>Domain</th>
<th>Time</th>
<th>DCT</th>
<th>DWT</th>
<th>WST</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPG2ABP</td>
<td>Input size</td>
<td>121 × 1</td>
<td>120 × 1</td>
<td>120 × 1</td>
<td>120 × 1</td>
</tr>
<tr>
<td></td>
<td>Output size</td>
<td>120 × 1</td>
<td>120 × 1</td>
<td>120 × 1</td>
<td>120 × 1</td>
</tr>
<tr>
<td>PPG2SBP/DBP</td>
<td>Input size</td>
<td>121 × 1</td>
<td>120 × 1</td>
<td>120 × 1</td>
<td>120 × 1</td>
</tr>
<tr>
<td></td>
<td>Output size</td>
<td>2 × 1</td>
<td>2 × 1</td>
<td>2 × 1</td>
<td>2 × 1</td>
</tr>
</tbody>
</table>
4.1 Beat-by-Beat cPPG-to-ABP mapping

The first scenario for estimating the BP is the estimation of the ABP which can be thought of as the continuous-time BP. In this section, the ABP beats are estimated from the corresponding PPG beats by using the LSTM network with a $120 \times 1$ sequence regressor output layer. As the output represents a sequence and we are interested in the time series ABP, there are seven possible training combinations for ABP estimation with different feature domains. These combinations are 1) using time domain (TD) for PPG beats with BI and TD for the ABP beats, 2) using DCT for PPG beats and TD for ABP, 3) using DCT for PPG beats and DCT for ABP and then using IDCT for comparison and evaluation, 4) using DWT for PPG beats and TD for ABP, 5) using DWT for PPG beats and DWT for ABP and then using IDWT for comparison and evaluation, 6) using WST for PPG beats and TD for ABP, 7) using WST for PPG beats and DWT for ABP and then using IDWT for comparison and evaluation. The last case is done with the WST-DWT combination because WST is not invertible and the DWT has the nearest features to the WST.

Table 3 shows the RMSE and MAE for the estimated ABP beats with different cases and different feature domains. As expected, using WST outperforms the other feature domains. Also, using the WST-DWT combination has a small enhancement compared to the WST-TD case. This is due to that learning the relation between the WST of PPG and DWT of the ABP is difficult and requires a complex network. Figure 8 shows an example of the reconstructed ABP beats with different feature domains and different cases compared to the ground truth ABP beat. As shown from this figure, the estimated ABP beat by using WST-TD and WST-DWT are highly related and correlated to the ground truth ABP beat.

<table>
<thead>
<tr>
<th>Domain</th>
<th>Time</th>
<th>DCT</th>
<th>DWT</th>
<th>WST</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case</td>
<td>TD + BI-TD</td>
<td>DCT-TD</td>
<td>DCT-DCT</td>
<td>DWT-TD</td>
</tr>
</tbody>
</table>

4.2 Beat-by-Beat cPPG-to-SBP/DBP mapping

Table 4 shows a comparison between different feature domains in sense of RMS error and MAE for the estimated SBP and DBP. From this table, it can be shown that using the time domain, DCT domain or DWT domain has a small effect on the BP estimation due to the sensitivity to the beat shift and scale. On the other hand, using WST has a recognized improvement of the RMSE and MAE due to better feature localization and insensitivity to shifting and scaling. Scattering plots of the predicted SBP and DBP versus the actual values are shown in Fig. 9. This figure is consistent with the RMSE and MAE values where the predicted SBP and DBP are less scattered in the case of WST domain compared to the time, DCT, and DWT domains.
4.3 Evaluation of the proposed method using rPPG signals

As the main goal of this work is to predict BP remotely by using a video camera, the proposed system is evaluated with the public rPPG data-sets [62] to show its performance with rPPG signals. In a similar way, the rPPG signals are segmented into beats and normalized to be in the range [0 − 1] using Eq. (3). The normalized beats are then normalized in time to be with the fixed length (120). The time interval is used as a feature besides the normalized beats for BP prediction.

As the case of cPPG, the performance of the predicted SBP and DBP by using different DL networks with the rPPG datasets. Figure 10 shows plots of three examples for the estimated ABP using Beat-By-Beat with different feature domain. From these figures it can be shown that using scattering wavelet transform as a feature domain can track the maximum value of ABP (SBP) as well as the minimum value of ABP (DBP). In addition, mean absolute error and mean square error for the estimated DBP and SBP from rPPG beats are shown in Table 5.

5. Discussion

Using LSTM time sequence-to-value regression, the SBP and DBP are estimated from the corresponding PPG features. The four feature domains are tested including the time domain, DCT domain, DWT domain, and WST domain. In the time domain, the beat interval information is included with the input sequence. Each feature domain has its advantages as follows: the time domain includes the beat interval and the behavior of the PPG which is related to the behavior of the ABP in the time domain and therefore the BP behavior. However, the time domain features require a huge dataset and complex network to extract the deep features directly from the PPG beats in the time domain. On the other hand, the DCT feature domain compresses the beats features in a small number of points which can help in reducing the input size with
less deformation. However, in this study, we used the full-length DCT features for a fair comparison. The main drawback of this feature domain is that error prediction of the DC and low-frequency components may lead to destructive results in the BP estimation and ABP prediction as well. The DWT domain has combinational features that are time and frequency which make it suitable for ABP and BP estimation. However, it suffers from the sensitivity to the signal shifting and scaling that usually happens with the PPG signals. Therefore, it may lead to errors due to scaling and shifting. Unlike DWT, WST doesn’t suffer from the effect of the shifting and scaling of the PPG beats. So, WST is the most suitable feature extractor that helps the LSTM network in learning the relationship between PPG and ABP and therefore BP estimation.

6. Conclusion And Future Work

Recently, BP is learned from ECG and/or PPG signals through machine learning or deep learning approaches. Here in this paper, we presented a beat-by-beat BP estimation from PPG beats only through a hybrid feature extraction and deep-learning approach. Two learning scenarios are introduced using different feature domains, namely, 1) Per-beat continuous PPG-to-ABP mapping, and 2) Per-beat PPG-to-SBP/DBP learning. The feature domain is introduced for helping the deep learning network in learning the complex relationships between PPG and ABP waveforms and resolving the limitations in the training dataset. Among the applied feature extractors, Wavelet scattering transform (WST) outperforms the other feature domains in sense of RMSE and MAE. Also, the reconstructed ABP beats are highly correlated to the ground truth ABP in the case of using WST-TD and WST-DWT. It is suggested as future work to 1) Replace LSTM with another deep learning network, and 2) Change the WST filtering based on another complex wavelet.

Declarations

Ethical Approval

Authors declare that the manuscript consists of public dataset for blood pressure and PPG signals and doesn’t include any human or animal studies.

Competing interests

Authors declare that the manuscript doesn’t include competing financial or personal interests.

Authors’ contributions

Authors’ contributions are as follows: Osama Omer, wrote the main manuscript text, implement the algorithm and figures preperation. Mostafa Salah shared in manuscript writing, algorithms implementation and figures preparation. Ammar M. Hassan share in manuscript preparation, results evaluation. Nirohiro Sugita shared in manuscript review and algorithm evaluation. Yoshifumi Saijo reviewed the manuscript.
Funding

Authors declare that they didn't receive any fund for this work.

References


**Figures**

**Figure 1**

Proposed Blood Pressure monitoring system (testing phase).
Figure 2

Beat segmentation based on local minimum detection.

Figure 3

Segmented beats (including noisy beats)

Figure 4
Selected beats (according to common selection criteria)

Figure 5

The block diagram of the proposed per-beat BP estimation system (training stage)

\[
S_0 x = x \star \phi
\]

\[
S x(t, \lambda_1) = |x \star \psi_{\lambda_1}| \star \phi
\]

\[
S x(t, \lambda_1, \lambda_2) = \|x \star \psi_{\lambda_1} \star \psi_{\lambda_2}\| \star \phi
\]

Figure 6
Multilayer wavelet scattering transform.

Figure 7

LSTM neural network architecture
Figure 8

Comparison between the estimated ABP beats from PPG beats using different transformations in two cases: a) High blood pressure, b) normal blood pressure, and c) low blood pressure.
Figure 9

the relationship between the predicted and actual SBP and DBP using: (a) Time domain featured, (b) Features in the DCT domain, (c) Features in the DWT domain, and (d) Features in the wavelet scattering domain.
Figure 10

the relationship between the predicted and actual SBP and DBP using: (a) Time domain featured, (b) Features in the DCT domain, (c) Features in the DWT domain, and (d) Features in the wavelet scattering domain.