Kahramanmaras Sutcu Imam University
Journal of Engineering Sciences

PHOTOPLETHYSMOGRAPHY BASED BLOOD PRESSURE ESTIMATION USING SYNCHROSQUEEZING TRANSFORM AND DEEP LEARNING

SENKRON SIKIŞTIRMA DÖNÜŞÜMÜ VE DERİN ÖĞRENME KULLANILARAK FOTOPLETİSMOGRAFİ TABANLI KAN BASINCI KESTİRİMİ

Yeşim HEKİM TANÇ[1] (ORCID: 0000-0001-8029-4253)
Mahmut ÖZTÜRK[1]* (ORCID: 0000-0003-2600-7051)

[1] Istanbul University - Cerrahpasa, Department of Electrical and Electronics Engineering, Istanbul, Türkiye

*Sorumlu Yazar / Corresponding Author: Mahmut ÖZTÜRK, mahmutoz@iuc.edu.tr

ABSTRACT

Cardiovascular diseases are one of the deadliest health problems. Hypertension is the most common reason for cardiovascular diseases. Keeping the blood pressure (BP) level under control is the only way to protect against the deadly results of hypertension. Therefore, monitoring BP regularly makes it possible to detect dangerous conditions in patients with hypertension. With the rapid developments in computers and sensor technologies, it is becoming possible to monitor BP levels continuously by using photoplethysmogram (PPG) signals. This work presents a non-invasive BP prediction method using one channel PPG signal. We employed the Synchrosqueezing Transform to obtain Time-Frequency (TF) images of the PPG signals. The TF images were used to feed a pre-trained deep neural network. We estimated the BP levels inside the 5-second intervals. Our method estimates BP levels with a mean error (ME) of 0.2148 mmHg and -0.0370 mmHg in the systolic and diastolic blood pressure (SBP and DBP) respectively. The ME values of our method are in the applicable levels. The standard deviation (SD) of our method is 5.0642 mmHg for DBP and 10.9904 mmHg for SBP. The upper limit specified by the AAMI is 8 mmHg. Also, our method is coherent with grades A and B according to the BHS standard.

Keywords: Photoplethysmography, arterial blood pressure, convolutional neural network, time-frequency analysis, synchrosqueezing transform.

ÖZET

Kalp ve damar hastalıkları en ölümcül sağlık sorunlarından biridir. Hipertansiyon, kardiyovasküler hastalıkların en yaygın nedenidir. Tansiyon düzeyini kontrol altında tutmak, hipertansiyonun ölümcül sonuçlarından kurtulmanın tek yoludur. Bu nedenle, hipertansiyonu olan hastalar için kan basıncının (KB) düzenli olarak izlenmesi, tehlikeli durumların tespitini mümkün kılar. Bilgisayar ve sensör teknolojilerindeki hızlı gelişmeler ile fotopletismogram (PPG) sinyalleri kullanılarak kan basıncının sürekli olarak izlenmesi mümkün hale gelmektedir. Bu çalışmada, tek kanallı PPG sinyali kullanan invazif olmayan bir KB tahmin yöntemi sunuyoruz. PPG sinyallerinin Zaman-Frekans (ZF) görüntülerini elde etmek için Eş Zamanlı Sıkıştırma Dönüşümünü kullandık. ZF görüntüleri, önceden eğitilmiş bir derin sinir ağını beslemek için kullanıldı. KB seviyelerini 5 saniyelik aralıklarla tahmin ettik. Yöntemimiz, sistolik ve diyastolik kan basıncı (SKB ve DKB) seviyelerini sırasıyla 0,2148 mmHg ve -0,0370 mmHg ortalama hata (ME) ile tahmin eder. Yöntemimizin ME değerleri uygulanabilir seviyelerdedir. Yöntemimizin standart sapması (SS) DKB için 5,0642 mmHg, SKB için 10,9904 mmHg değerindedir. AAMI tarafından belirlenmiş olan üst limit 8 mmHg değerindedir. Ayrıca yöntemimiz BHS standardına göre A ve B sınıfları ile uyumludur.

Anahtar Kelimeler: Fotopletismografi, arteriyel kan basıncı, konvolüsyonel sinir ağı, zaman-frekans analizi, senkron sıkıştırma dönüşümü.

INTRODUCTION

Hypertension is a very common and fatal health problem for human beings in modern life. It has been guessed that approximately 1.25 billion people have tension problems all over the world (World Health Organization, 2022). The reason for hypertension has been determined as blood pressure (BP) abnormalities by researchers decades ago. When having this kind of BP problem, tracing the BP level continuously and keeping it under control is very crucial. Classically, medics use the cuffed sphygmomanometer to measure the blood pressure (BP) level. The accuracy of the cuffed measurements changes depending on the ability and experience of the medics. Moreover, BP level monitoring using cuffed measurements is not a practically applicable method in patients' daily life. To overcome such obstacles, photoplethysmography (PPG) signals can be used to monitor and trace a person's blood pressure continuously in daily life.

Blood pressure can be measured using invasive or non-invasive methods. It is clear that the invasive measurement provides more accurate measurement results than the noninvasive methods. Unfortunately, invasive measurement can be applied only in hospitals or clinics. In daily life, noninvasive measurement of BP is more applicable. The traditional cuff sphygmomanometer is the most used noninvasive BP measurement method. Although it is a noninvasive method, it is not practical to use often in daily life. Some experience is necessary to use these devices correctly. Also, it is not comfortable for patients who need to frequently monitor and trace their BP levels. PPG based BP estimation is a noninvasive and cuffless measurement method. The usage of PPG signals for BP monitoring is easier, more practical, and comfortable for the patients. With the developing recording devices and increasing accuracy rates, this method has become more popular in the last decades. Noninvasive, easier and practical measurement is possible using PPG sensors, and these properties make it useful in daily routine for frequent monitoring of BP levels. A smartwatch, a wristband, or the devices used in hospitals can be designed to record PPG signals and predict BP levels.

Recent developments in computer technologies and sensors make it possible to obtain some biological values of people noninvasively. For real-time monitoring and tracing, PPG sensors have been used widely. By using PPG sensors, comfortable, noninvasive, and cuffless monitoring of many biological values is possible. PPG signals are recorded using optical sensors. Some researchers used PPG signals to estimate heart rate and oxygen saturation levels; in this way, they made important contributions to diagnosing respiratory problems and sleep apnea (Oğuz et al., 2023; El-Hajj and Kyriacou, 2020; McDuff et al., 2014; Salehizadeh et al., 2014; Lázaro et al., 2014; Yousef et al., 2012; Shin et al., 2009; Suzuki et al., 2009; Arnold et al., 2007; Johnston, 2006; Kim and Yoo, 2006; Kraitl and Ewald, 2005). Also, a growing number of research have been focused on the estimation of BP using PPG signals (El-Hajj and Kyriacou, 2020).

Cuffless BP estimation is a challenging research area. Because of the importance of BP level tracing for health conditions, PPG based BP estimation research has been getting more popular since the last decade. Some promising methods have been proposed by researchers in the last years and high accuracies have been obtained using machine learning methods. Some researchers preferred to use shape-based time-domain features to estimate BP levels (Chao et al., 2021). The preferred signal features for PPG based BP estimation used in recent studies have caused low estimation errors and high accuracies. However, their recording processes are not easy and comfortable; also the necessary computations for extracting them are rather complex. Some of the features introduced and used in the recent BP estimation studies are pulse transit time (PTT), signal amplification time, pulse arrival time (PAT), pulse wave velocity (PWV), PPG pulse shape, systolic upstroke time (SysT), diastolic time (DiasT), pulse width (PW), and pulse height (Chao et al., 2021). All of these features depend on the shape of PPG pulses. Therefore, they are very sensitive to noise and patient movements. Even some small changes in the time domain cause significant differences in the estimation results. This sensitivity is not affordable in real-time tracing applications. Moreover, for extracting PTT two different but simultaneous PPG signals of different body parts are necessary. For calculating the PAT an additional but simultaneous ECG signal must be recorded.

Despite the high sensitivity to noise and small synchronization errors in the time domain, some researchers managed to get good estimation accuracies by using shape-based features. Liu et al. proposed a new method for real-time and continuous BP estimation using synchronous ECG and PPG signals (Liu et al., 2020). They preferred versions of PTT, PAT, PW, SysT, and DiasT values as features extracted from the PPG and ECG signals. With the help of machine learning (ML) methods, they estimated systolic blood pressure (SBP) and diastolic blood pressure (DBP) separately. Pour Ebrahim et al. presented an algorithm for measuring SBP continuously (Pour Ebrahim et al., 2019).

KSÜ Mühendislik Bilimleri Dergisi, 27(1), 2024
Araştırma Makalesi

245

KSU J Eng Sci, 27(1), 2024
Research Article

Y. Hekim Tanç, M. Öztürk

They used simultaneously recorded ECG and PPG signals to extract shape-based features. Teng et al. proposed a method for BP estimation using only one channel PPG recording (Teng et al., 2003). They also used different versions of shape-based features and regression algorithms. Ruiz-Rodriguez et al. monitored BP levels continuously using one channel PPG (Ruiz-Rodriguez et al., 2013). Xing and Sun suggested using frequency-domain features (Xing and Sun, 2016). They used classical Fast Fourier Transformation (FFT) to obtain the spectrums of PPG signals. They extracted spectrum-based features to use the ML stage.

The usage of deep learning (DL) methods for BP estimation has become more popular in the last few years. Sun et al. used pre-trained convolutional neural networks (CNN) to predict BP from PPG data (Sun et al., 2021). They used Hilbert-Huang Transforms (HHT) of PPG signals and their first and second derivatives. They obtained the best classification accuracies with AlexNet. Liang et al. proposed a method using Continuous Wavelet Transform (CWT) scalograms of PPG signals and a pre-trained CNN (Liang et al., 2018). They preferred GoogLeNet and their classification accuracies were between 80% to 90% in different scenarios. Tjahjadi et al. presented a new approach using Bidirectional Long-Short Term Memory (BLSTM) at DL stage (Tjahjadi et al., 2020). They used the Short-Time Fourier Transform (STFT) of PPG signals to estimate BP levels. Classification accuracy of their approach was higher than 93%. El-Hajj and Kyriacou used raw PPG signals and their derivatives for feeding CNN structure (El-Hajj and Kyriacou, 2021). In the studies presented by (Tazarv and Levorato, 2021) and (Esmaelpoor et al., 2020), the authors preferred to use different versions of LSTM based CNN structures to predict BP levels of patients.

SST is a very robust method against noise and time-domain changes. Also, its time and frequency resolutions are very high compared to other TF analysis methods. The PPG signals were divided into 5-second segments, and SST was applied to each segment separately. At last, a single image of SST representation is obtained and stored for each 5-seconds segment. Then, the SST representation images are used for feeding the pre-trained CNN. As a result, we have estimated the SBP and DBP values for each of the 5-second segments and have shown that the continuous tracing of BP levels using PPG signals is possible.

The contributions of this study can be summarized as 1) Our approach uses only one channel PPG recording for BP estimations, 2) PPG records can be obtained very easily compared to traditional cuff-based BP measurement, 3) The proposed method is robust because the shape-based features are not used, 4) The Synchrosqueezing Transform (SST) which we preferred to use while obtaining TF representations of PPG signals provides very high time and frequency concentrations.

DATASET

In this study, the Multiparameter Intelligent Monitoring in Intensive Care (MIMIC-II) database is used to obtain the arterial blood pressure (ABP) and PPG signals. This database has been formed with the physiological signal records of intensive care unit patients (Saeed et al., 2011). It contains more than 23,000 records. Duration of the records changes between seconds to several hours; sampling frequencies are 125 Hz; and precision is at least 8-bit. The MIMIC-II dataset contains waveforms of ECG, PPG, and respiration, some numeric data such as heart and respiration rates, SpO_2, and systolic, mean, and diastolic blood pressures. PPG signals were obtained using a pulse oximeter at the finger. BP levels were measured invasively at the wrist. In our study, we used a preprocessed, filtered, and diminished version of the MIMIC-II dataset formed by Kachuee et al. (Kachuee et al., 2017). MIMIC-II dataset originally include many recordings with varying lengths and signals. Some records do not include PPG signals, some of them do not include blood pressure levels. In some records, signals are very short for meaningful analysis. Therefore, it is not possible to use all records in this research. For this reason, we decided to use a cleared, pre-processed, and reduced version of the MIMIC-II dataset. The dataset of Kacuhee et al. is more useful for this kind of research (Kachuee et al., 2017). At last, we took 5599 records from the database and divided them into 12 segments. The signal records included in the database are in a cell array of matrices form. Each cell includes a single record and each row of the matrices shows one signal channel, i.e. PPG, ABP, and ECG from channel II respectively (Silva et al., 2014). A sample of signals in a 5 second-segment can be seen in Fig.1.

METHODOLOGY

Synchrosqueezing Transform

Traditionally, spectrums of the signals are analyzed using the Fourier Transform (FT) and some similar methods. They are useful tools for obtaining the frequency contents of the signals. The absolute square of FT forms the distribution of the signal's power depending on the frequency content. However, this analysis method does not give

information about which frequencies exist in which time periods. Therefore, FT and similar approaches are not satisfying to analyze the signals that have time-varying spectrums. Signals in real life problems are mostly random signals, especially biomedical ones. For obtaining and understanding the frequency contents and time-varying power spectrums of the random signals, Time-Frequency (TF) analysis methods can be used.
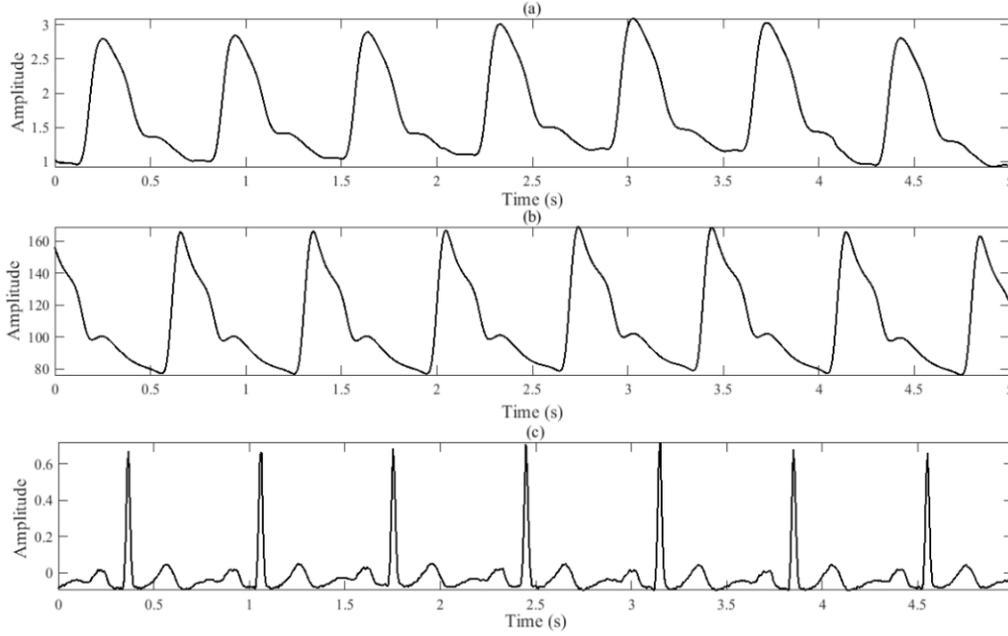
Figure 1. A 5 Second Segment of (a) PPG Signal, (b) ABP Signal, and (c) ECG Signal

Short-Time Fourier Transform (STFT), Wigner-Ville Distribution (WVD), and Wavelet Transform (WT) are the most known and widely used TF analysis methods in the last decades. Also, different versions of these fundamental methods have been developed for obtaining TF representations with higher time and frequency resolutions. Synchrosqueezing Transform (SST) is a rather new and promising TF analysis method.

In this work, we use SST to obtain time-frequency representations of PPG signals. SST provides an excellent resolution in time-frequency representations. It takes one of the popularly known time-frequency methods and improves it. So, SST can be classified as a Reassignment Method (RM). In our approach, we preferred to use Wavelet based SST while analyzing the PPG signals. In short, SST estimates the Instantaneous Frequency (IF) of the signal with the help of Continuous Wavelet Transform (CWT). By using IF information, SST algorithm sums the spread energy of the signal and concentrates it through the IF. This part of the algorithm can be thought of as the squeezing of energy around the IF. As a result, a highly concentrated time-frequency representation is obtained. It can be seen in several researches (Auger et al., 2013; Daubechies et al., 2011) that representation accuracy of SST is much higher than traditional TF techniques.

The main advantage of SST is the high time and frequency resolutions in TF representations as mentioned before. Better representation of the signals provides better results in signal processing based research. Applications like noise filtering, artifact reduction, etc. become easier and more successful. Also, the performances of classification and regression algorithms increase with the usage of highly concentrated TF representations. These advantages of SST exist for all kinds of signals not only for PPG signals. Its success and advantage compared to the other TF methods can be seen especially in analyzing the random signals. PPG is also random like the other biomedical signals. Obtaining the TF representations of random signals with high resolutions is a challenging procedure using traditional signal analysis techniques. SST provides highly concentrated TF representations for random signals. Therefore, it is applicable to all kinds of biomedical signals like PPG signals.

SST has been developed by Daubechies et al. (Daubechies et al., 2011). It can be thought like a post-processing method used with CWT. First of all, the CWT of the signal $x(t)$ is calculated:

$$W_x(a,b) = \int_{-\infty}^{\infty} x(t) a^{-1/2} \bar{\psi}\left(\frac{t-b}{a}\right) dt \qquad (1)$$

Here, $W_x(a, b)$ represents the wavelet coefficients for each pair of (a, b). a shows the scale factor that is inversely proportional to the frequency. b is the time-dependent translation factor. $\bar{\psi}(x)$ represents the complex conjugate of the mother wavelet $\psi(x)$.

The IF of the signal can be estimated using the wavelet coefficients $W_x(a, b)$, as follows for any (a, b) where $W_x(a, b) \neq 0$ (Daubechies et al., 2011):

$$\omega_x(a, b) = \frac{-j}{2\pi W_x(a, b)} \frac{\partial W_x(a, b)}{\partial b} \tag{2}$$

With the help of IF, all wavelet coefficients $W_x(a, b)$, are reallocated into $T_x(\omega_l, b)$. Here, ω_l is the closest frequency to the real IF, $\omega_x(a, b)$. Finally, SST is obtained as follows (Daubechies et al., 2011):

$$T_x(\omega_l, b) = (\Delta\omega)^{-1} \sum_{a_k:|\omega_x(a_k,b)-\omega_l|\leq\frac{\Delta\omega}{2}} W_x(a_k, b) a_k^{-\frac{3}{2}} (\Delta a)_k \tag{3}$$
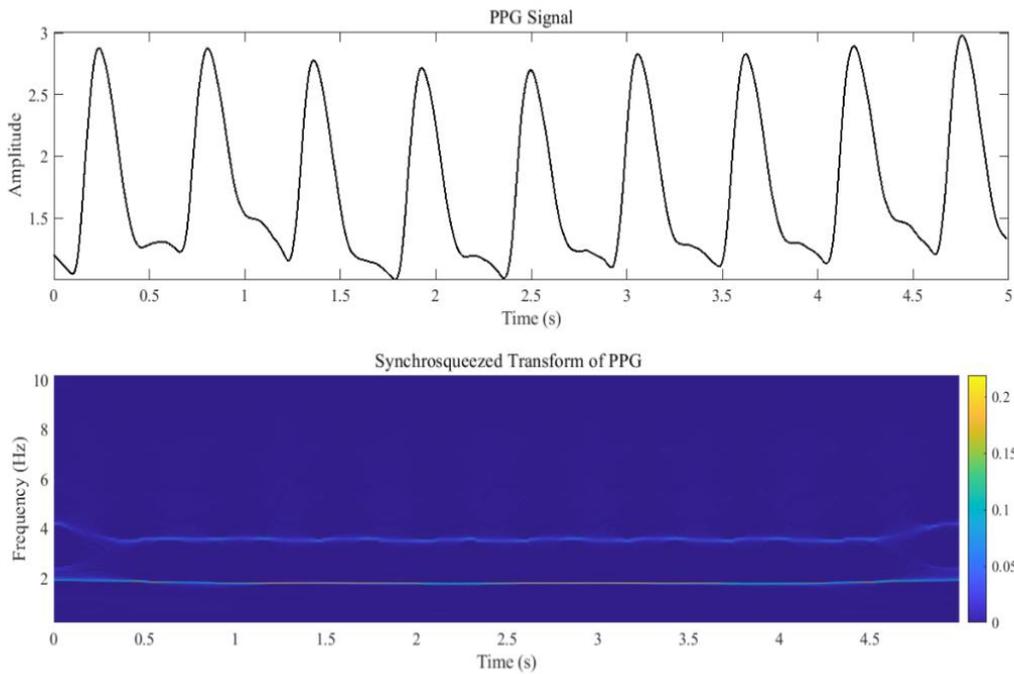
Figure 2. A 5-Second Segment of PPG Signal and its SST Spectrum

Here, $T_x(\omega_l, b)$ shows the SST coefficients at the centers ω_l of consecutive frequency bins; $\Delta\omega$ is the width of these frequency bins, $[\omega_l - \frac{1}{2}\Delta\omega, \omega_l + \frac{1}{2}\Delta\omega]$, $\Delta\omega = \omega_l - \omega_{l-1}$, and $(\Delta a)_k = a_k - a_{k-1}$. Briefly, for each time point b, the reassignment frequencies are obtained using (2) for all scales. After that, the SST coefficients $T_x(\omega_l, b)$, have been calculated as the sum of CWT coefficients, $W_x(a, b)$, where the difference between the reassigned frequency $\omega_x(a, b)$ and ω_l is less than a specified bin width $\frac{\Delta\omega}{2}$.

Time-Frequency Analysis of PPG Signals

At the experimental stage of this study, 100 data records of patients with 9 minutes and longer lengths ABP and PPG signals from the modified MIMIC-II database were selected randomly. Firstly, the signals were divided into 5s segments. Then, we filtered each PPG signal segment with a 0.5–10 Hz Chebyshev Type II filter. This bandwidth is chosen for bandpass filter because the change rate of the PPG signals is directly related to the heartbeat. Therefore, the frequencies of the PPG signals are expected lower than 4 Hz. SST was applied to each segment separately and TF representations of all segments were obtained as can be seen in Figure 2. It is seen that there is no signal component above approximately 8 Hz. Therefore, we filtered the frequencies higher than 8 Hz before forming TF images of the PPG segments. After that, all TF images were resized to feed the convolutional neural network. We

| KSÜ Mühendislik Bilimleri Dergisi, 27(1), 2024 | 248 | KSU J Eng Sci, 27(1), 2024 |
| Araştırma Makalesi | | Research Article |

Y. Hekim Tanç, M. Öztürk

preferred to use ResNet-18 for the solution of the regression problem, so the sizes of TF images of the PPG segment had to be $224 \times 224 \times 3$. Then, the images of each segment were combined to form a four-dimensional (4-D) array.

The Arterial Blood Pressure (ABP) signals obtained from the dataset can be used as the ground truth for SBP and DBP values. The ABP signals show the results of periodical BP measurements. Classical techniques were employed for these measurements. For extracting the ground truth values of SBP and DBP from ABP, the method proposed by Elgendi et al. was used (Elgendi et al., 2013). In this method, SBP is the highest value in the segment of heart's systole, and DBP is the lowest value of the valleys in the segment of heart's diastole.
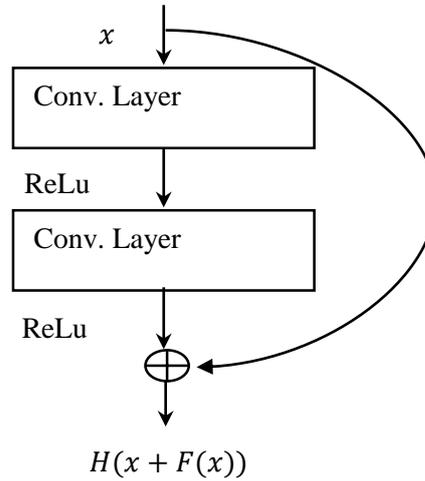
Figure 3. Architecture of a Residual Block

ResNet-18 Architecture for Regression Analysis

The Residual Network (ResNet) architecture proposed by He et al. is one of the most significant CNN models (He et al., 2016). It manages to classify many images quickly with high accuracy. With the insertion of residual blocks, ResNet architecture protects itself from the vanishing gradient problem and degradation which are known as common problems in CNN based structures and occur as an expected result of depth increase. Jumping some connections in residual blocks cancels the effects of the jumped layers in the neural network and transmits the output of a previous layer as the input to the next layers. The standard rectified linear activation (ReLu) function is $H(x) = ReLu(x) = \max(0, x)$. Architecture of a basic residual block is illustrated in Figure 3.

Many different ResNet architectures have been proposed and used by researchers (Sunnetci et al., 2023). We preferred to use ResNet-18 in this work. Its depth and computational load seem to be enough and efficient for this kind of regression analysis. ResNet-18 includes 5 convolution blocks, 17 convolution layers, and 1 fully connected layer. The first convolution block can be thought of as a single convolution layer with 64 (7×7) filters. The following convolution blocks are formed with two residual blocks. Each residual block includes 2 convolution layers and these layers have same number of 3×3 filters. Also, each convolution block decreases the image size by half while increasing the feature dimension by two. Different from the standard CNN structure, each 3×3 filter pair has a shortcut connection.

ResNet-18, a variant of the Residual Network (ResNet) architecture, is a popular choice in deep learning and computer vision tasks, including image classification, due to several advantages that make it a strong candidate for a wide range of problems. We can elaborate on why ResNet-18 could be a suitable choice and its advantages over other potential architectures, taking into consideration the nature of our data:

Depth and Computational Load: ResNet-18 strikes a balance between model depth and computational load. It is not as deep as some other architectures like ResNet-50 or VGG-16, which can be computationally expensive and memory-intensive. For certain applications where computational resources are limited, ResNet-18 may be more practical without sacrificing too much performance.

Skip Connections and Residual Blocks: The key innovation in ResNet is the use of skip connections and residual blocks. These skip connections allow gradients to flow more easily during training, mitigating the vanishing

gradient problem. This is especially beneficial when dealing with deep networks, which is often the case in computer vision tasks. The presence of skip connections can help in capturing fine-grained details in our data.

Efficiency in Training: Architecture of ResNet helps in training deeper networks more effectively. This can be particularly advantageous if our dataset is large and complex. It enables the model to learn both low-level and high-level features efficiently, which is often crucial in image classification tasks.

Transfer Learning: ResNet-18 is a well-established deep learning architecture that has been pre-trained on large datasets like ImageNet. This pre-training allows us to boost transfer learning. We can start with a pre-trained ResNet-18 model and fine-tune it on our specific dataset. Transfer learning is valuable when we have limited labeled data because it helps the model to generalize better from the pre-trained knowledge.

Regularization: The skip connections of ResNet act as a form of regularization. They help to prevent overfitting, which is important when dealing with limited data. This can lead to a better generalization performance on our specific dataset.

Community and Resources: ResNet-18 has been widely adopted in the deep learning community, and there are plenty of resources, pre-trained models, and fine-tuning strategies available. This can save us time and effort in model development and training.

Tested Performance: ResNet-18 has demonstrated strong performance on a variety of computer vision tasks, including image classification, object detection, and segmentation. Its effectiveness in these tasks suggests that it is a reliable choice for a wide range of problems.

In summary, the advantages of ResNet-18, such as its efficient training, regularization properties, and strong performance on various computer vision tasks, make it a solid choice for image classification problems, especially when we have limited data or computational resources. While other architectures may have their strengths, ResNet-18's balance of depth and performance often makes it a sensible starting point for many real-world applications in computer vision, including those with unique data characteristics.

Because the ResNet-18 structure was designed originally for classification problems, we made some changes on the architecture to use it in a regression problem. We changed the final classification layer of the ResNet-18 with a regression output layer and a fully connected layer of size 1 that relates to the number of responses BP levels. The final structure after our modifications was composed of 71 layers. Moreover, we increased the learning rate factor for the weights and the biases of the fully connected layer to obtain faster learning in the new layers than in the transferred layers. The ResNet-18 architecture used in this study can be seen in Figure 4.

Adapting the ResNet-18 architecture for regression analysis typically involves specific modifications to the output layer and the loss function used for training. Here's how ResNet-18 can be adapted for regression, including the changes to the output layer and the loss function:

In a standard ResNet-18 architecture for image classification, the output layer typically consists of a fully connected layer followed by a softmax activation function, which is suitable for multi-class classification tasks. However, for regression analysis, where the goal is to predict a continuous numerical value, this output layer needs to be modified. For regression, we would replace the final fully connected layer with a single neuron or a small fully connected layer without any activation function. The output from this layer would be a continuous numerical value, which is the regression prediction. In the case of a single neuron output, this value represents the predicted target or continuous variable we are trying to estimate.

The choice of loss function is critical in regression tasks. Instead of using the categorical cross-entropy loss commonly used in classification tasks, we would typically use a regression-specific loss function. The most used loss function for regression is the Mean Squared Error (MSE) loss. The goal during training is to minimize this MSE loss. It quantifies how well the model's predictions match the actual target values. With the modified output layer and the MSE loss function, we can train our ResNet-18 architecture for regression. During training, the model adjusts its weights and biases to minimize the MSE loss, which means it is trying to make its predictions as close as possible to the real target values in our dataset. We would typically use gradient-based optimization algorithms, such as

| KSÜ Mühendislik Bilimleri Dergisi, 27(1), 2024 | 250 | KSU J Eng Sci, 27(1), 2024 |
| Araştırma Makalesi | | Research Article |

Y. Hekim Tanç, M. Öztürk

stochastic gradient descent (SGD) or its variants (e.g., Adam), to update the model's parameters and minimize the loss function.

As a result, when adapting ResNet-18 for regression analysis, we modify the output layer to have a single unit with a linear activation function and choose an appropriate loss function like MSE or MAE. These modifications allow the network to predict continuous values and optimize its parameters based on the error between predictions and real target values. These changes are fundamental in transforming a classification-focused neural network architecture into one suitable for regression tasks.
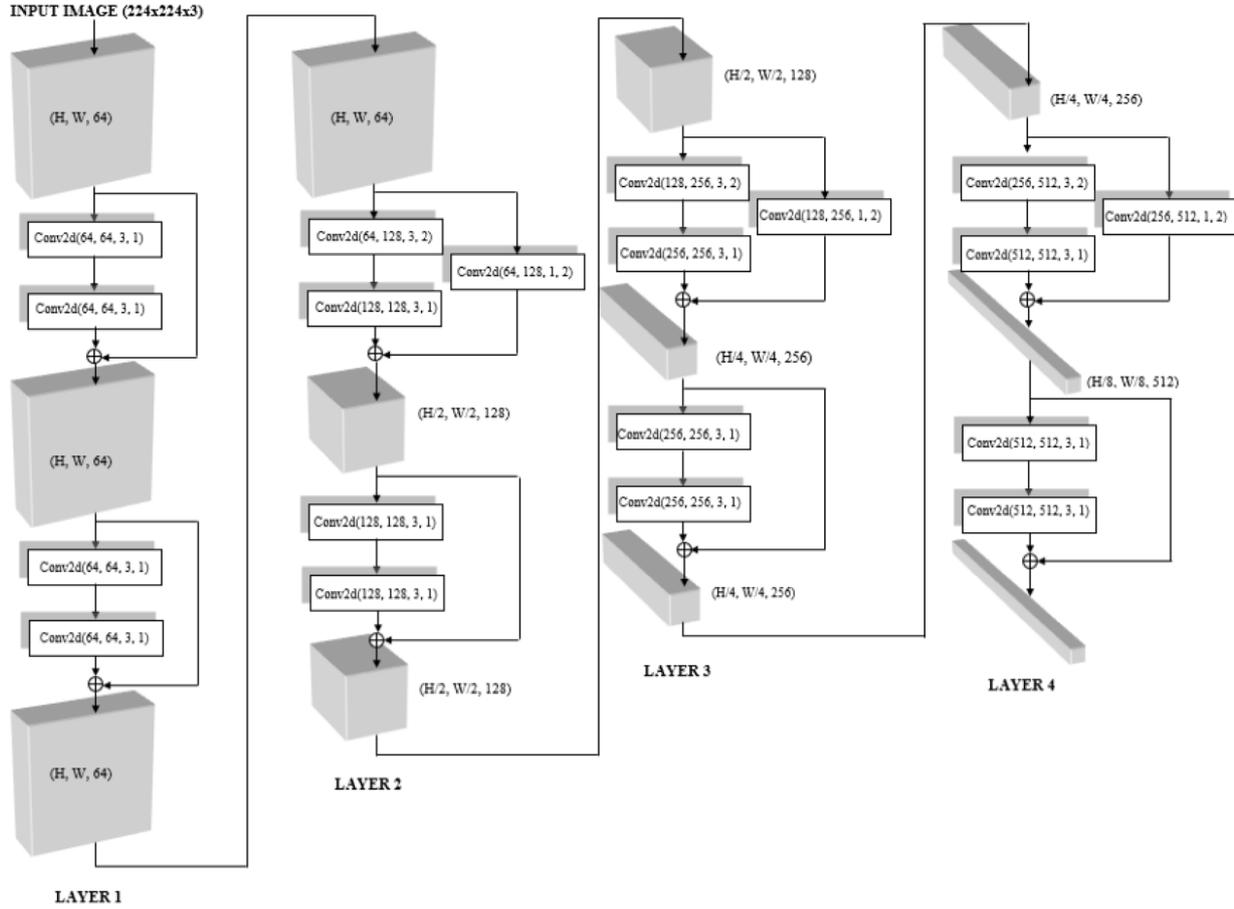
Figure 4. Architecture and Parameters of ResNet-18

RESULTS AND DISCUSSIONS

In this research, we used the pre-trained deep learning network ResNet-18. BP levels were estimated by regression analysis on ResNet-18 architecture. TF images obtained using SST of the PPG signals were used in the training and testing stages of the deep learning network. ADAM optimizer was used in the training of the network with a batch size of 128 and 10^{-4} learning rate. We limited the number of training epochs to 25. The learning rate factors for the weights and the fully connected layer's biases were set to 10. 70% of the TF images were used for training and 30% of them were used for validation. Training and validation sets were specified randomly. All stages of the algorithm were performed on MATLAB software (version R2021b).

For measuring the successes of predictions, we used the mean error (ME), mean absolute error (MAE) and the standard deviation (STD) of the error. The performance metrics are below (Lang et al., 2022; Kayadelen et al., 2022):

$$ME = \frac{1}{n}\sum_{i=1}^{n}\left(BP_{predicted}^{i} - BP_{ground\ truth}^{i}\right) \tag{4}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|BP_{predicted}^{i} - BP_{ground\ truth}^{i}\right| \tag{5}$$

$$STD = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(BP_{predicted}^{i} - BP_{ground\ truth}^{i} - ME\right)^2} \qquad (6)$$

Here n is the number of estimated BP levels for validation data, $BP_{predicted}$ expresses the predicted BP value and $BP_{ground\ truth}$ expresses the reference BP value.

According to the Association for the Advancement of Medical Instrumentation (AAMI), as a standard for reliable BP measurement, the maximum ME has to be 5 mmHg, and the maximum STD has to be 8 mmHg (American National Standards Institute, 2017). Our results can be seen in Table 1 for systolic blood pressure (SBP) and diastolic blood pressure (DBP) separately. ME values in the proposed method are lower than the maximum allowed ME value. It is seen that our approach is mostly successful, but only the STD of SBP estimations is higher than the AAMI permissible limit. We believe that the reason for this high STD value of our estimations is due to the high STD of SBP recordings in the dataset. SBP values in the dataset spread to a wide range and this makes the regression analysis more challenging.

Table 1. Validation Results According to AAMI Standard

		ME (mmHg)	STD (mmHg)
AAMI	SBP	5	8
	DBP	5	8
Our Results	SBP	0.2148	10.9904
	DBP	-0.0370	5.0642

According to the British Hypertension Society (BHS) (O'Brien et al., 2001), a BP measurement device must have a minimum BP Monitor Grade of B for utilization in a clinical setting. The grade is the cumulative proportion of the prediction error that falls within 5, 10, and 15 mmHg of the mercury standard. The validation results according to the BHS standard can be seen in Table 2. Our model is at the grade A in DBP estimation and at the grade B in SBP estimation.

Table 2. Validation Results According to BHS Standard

		$\leq 5\ mmHg$	$\leq 10\ mmHg$	$\leq 15\ mmHg$	Class
BHS		60%	85%	95%	Grade A
		50%	75%	90%	Grade B
		40%	65%	85%	Grade C
			Worse than C		Grade D
Our Results	SBP	74.97%	87.06%	92.04%	Grade B
	DBP	88.73%	97.28%	99.03%	Grade A

In Table 3, the comparison of our work with some other PPG based BP estimation methods has been presented. Only the methods that used regression were taken for comparisons. Although we have used only one channel of PPG signals, the performance of our method is better than most of the other methods that used two channels of PPG, PPG, and ECG together or modified versions of PPG. According to the performance metrics, the results of our approach are remarkable and promising for the design of real-time BP monitoring devices.

The studies that have used PTT, PAT, and similar waveform–based features are generally weak against time-domain manipulations and noise. Moreover, they need to use a second PPG recording or an additional ECG recording. These necessities make it difficult to use these methods in real life applications. Recording a second signal is not comfortable. Also, because of the movements of an individual in real life, noise, and time-shiftings can be realized. These reasons decrease the estimation performances. As seen in Table 3, our approach has lower error rates than the methods which use waveform-based features.

The estimation methods that use deep learning have lower error rates than our method in SPB estimation. In DBP estimation, the best results are obtained by Esmaelpoor et al. (Esmaelpoor et al., 2020). Our results are a little bit higher than the other methods. The difference is not much and significant. Our results are better than the results of Slapnicar et al. (Slapnicar et al., 2019) and Rodriguez et al. (Rodriguez et al., 2013). These differences can be raised as a result of the different signal processing techniques and deep learning structures.

Table 3. Performance Comparison with Related Works

Methods	Signals & Methods	Performance Metrics	DBP	SBP
(Kachuee et al., 2017)	PTT features	MAE ± STD	6.34 ± 8.45	12.38 ± 16.17
(Kachuee et al., 2015)	PAT features	MAE ± STD	5.35 ± 6.14	11.17 ± 10.09
(Tanveer and Hasan, 2019)	PTT raw signals	MAE	0.52	0.93
(Kurylyak et al., 2013)	PTT raw signals	ME ±STD	2.21 ± 2.09	3.80 ± 3.46
(Slapnicar et al., 2019)	PPG derivatives and raw signals + CNN	MAE	12.38	15.41
(El-Hajj and Kyriacou, 2021)	PPG features + CNN	MAE ± STD	2.6 ± 4.41	4.51 ± 7.81
(Rodriguez et al., 2013)	PPG waveform features + CNN	ME ±STD	-3.65 ± 8.69	-2.9 ± 19.35
(Tazarv and Levorato, 2021)	CNN-LSTM	MAE ± STD	2.02 ± 1.76	3.70 ± 3.07
(Esmaelpoor et al., 2020)	CNN-LSTM	MAE ± STD	0.67 ± 2.84	1.91 ± 5.55
Proposed method	PPG raw signals + ResNet-18	ME ±STD / MAE ±STD	-0.03 ± 5.06 / 3.51 ± 5.06	0.21 ± 10.99 / 7.52 ± 10.99

CONCLUSIONS

In this research, we proposed a method to estimate SBP and DBP values using only one channel PPG recordings. In contrast with the other research on this problem, neither ECG records nor shape-based features were used to extract features in time or frequency domains. The high-resoluted TF images were generated as SST of PPG signals and they were applied to a deep learning algorithm for regression analysis. We preferred to use the ResNet-18 in regression analysis which is a pre-trained deep learning architecture. The reason for this choice is because of its efficiency and speed in image classification problems. Moreover, we modified the ResNet-18 for using it in regression analysis. The speed and computational efficiency of the pre-trained Resnet-18 architecture make it promising for being used in real-time BP monitoring systems.

Our algorithm is robust against noise and motion artifacts owing to the advantages of working in joint TF domain and excellent representation capability of SST. Reassignment based algorithm of SST makes the TF representations robust. Especially using a real-time BP monitoring device, our algorithm has some advantages against the challenges of real-time monitoring. The main problems are different kinds of noise and motion artifacts in real-time BP estimation. The reason for these problems is the movements of subjects. The usage of SST makes it easier to filter noise and artifacts because its algorithm concentrates the signals through the high energy regions or instantaneous frequencies of them. This squeezing algorithm deletes most of the noise and artifacts.

Besides low error rates and robustness, usage of only one channel PPG signal makes our method promising for the design of a real-time BP monitoring system. This algorithm can be used in wearable devices like smart watches and wristbands to monitor BP in daily life. Also, PPG signals can be recorded from the fingers and the earlobes of the patients easily in the hospital environment for monitoring the BP continuously.

The main limitation of our approach is the high computational loads of SST and deep learning, but luckily, the developments in the capabilities of processors will eliminate the computational loads of them in the near future. The importance of this limitation increases especially in real-time applications. Nowadays, it is possible to implement this algorithm using a fast processor and high capacity memory. However, it is not efficient because of the high prices of these hardwares. Another limitation of our method is the usage of only one channel PPG signal. A second PPG signal or an ECG signal makes it possible to obtain estimations with lower error rates. However, one channel PPG signal based BP estimation is more suitable and preferable for applications in daily life.

As a future research perspective, we aim to develop an optimized algorithm with efficient computations. The high computational load bil of our method can be reduced by using different signal processing techniques and deep learning structures. In signal processing side, low computations reduce also the resolutions of TF representations. This causes a possible performance decrease in regression analysis. A computationally efficient TF method with ignorable performance decrease can still be found and employed. On the other side, different deep learning architectures and modified versions of them can be tried to reduce the load of the learning stage. Another option is the usage of machine

learning algorithms instead of deep learning. Machine learning methods need less computation than deep learning methods.

In this study, we used only the modified version of MIMIC-II dataset. This dataset has been used in many different research and it is still useful. However, the importance of testing the method with different datasets can not be ignored. In this stage, we aim to improve the performance of our method by employing some other signal processing techniques and learning algorithms. After obtaining better prediction results, we are going to test our approach with different datasets and in a real life application.

ACKNOWLEDGEMENTS

REFERENCES

American National Standards Institute. ANSI/AAMI/ISO 81060–2:2013.Non-invasive sphygmomanometers - Part 2: clinical investigation of automated measurement type. http://webstore.ansi.org, Accessed July 15, 2017.

Arnold, M. A., Liu, L.,& Olesberg, J. T., (2007), Selectivity assessment of noninvasive glucose measurements based on analysis of multivariate calibration vectors, *Journal of Diabetes Science and Technology*, 1(4), 454–462.

Auger F., Flandrin P., Lin Y., et al, (2013). Time-Frequency reassignment and synchrosqueezing, *IEEE Signal Processing Magazine*, 30:32-41.

Chao, P. C. , Wu, P. C. -C. D., Nguyen, H. B. -S. Nguyen, P. -C. Huang and V. -H. Le, (2021). The Machine Learnings Leading the Cuffless PPG Blood Pressure Sensors Into the Next Stage, *IEEE Sensors Journal*, 21(11), 12498-12510, doi: 10.1109/JSEN.2021.3073850.

Daubechies, I., Lu, J., Wu, H.T., (2011), Synchrosqueezed wavelet transform: an empirical mode decomposition like tool, *Appl. Comput. Harmon. Anal*., 30(2), 243-261.

Elgendi M, Norton I, Brearley M, Abbott D, Schuurmans D. (2013), Systolic peak detection in acceleration photoplethysmograms measured from emergency responders in tropical conditions. *PLoS One*, 8:e76585.

El-Hajj, C., and Kyriacou, P.A, (2021), Cuffless blood pressure estimation from PPG signals and its derivatives using deep learning models, *Biomedical Signal Processing and Control*, 70, 102984.

El-Hajj, C., and Kyriacou, P.A., (2020), A review of machine learning techniques in photoplethysmography for the non-invasive cuff-less measurement of blood pressure, *Biomedical Signal Processing and Control*, 58,.

Esmaelpoor, J., Moradi, M. H., and Kadkhodamohammadi, A., (2020), A multistage deep neural network model for blood pressure estimation using photoplethysmogram signals, *Computer in Biology and Medicine*, 120, 103719.

He, K., Zhang, X., Ren, S., and Sun, J., (2016), Deep Residual Learning for Image Recognition, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770-778, doi: 10.1109/CVPR.2016.90.

Johnston, William S., (2006), Development of a Signal Processing Library for Extraction of SpO2, HR, HRV, and RR from PhotoplethysmographicWaveforms, Masters Theses (All Theses, All Years), 919, https://digitalcommons.wpi.edu/etd-theses/919.

Kachuee, M., Kiani, M. M., Mohammadzade, H., and Shabany, M., (2017), Cuffless Blood Pressure Estimation Algorithms for Continuous Health-Care Monitoring, *IEEE Transactions on Biomedical Engineering*, 64,(4), 859-869.

Kachuee, M., Kiani, M. M., Mohammadzade, H., and Shabany, M., (2015), Cuff-less high-accuracy calibration-free blood pressure estimation using pulse transit time, 2015 IEEE International Symposium on Circuits and Systems (ISCAS), pp. 1006-1009.

Kayadelen, C., Altay, G., Önal, S., & Önal, Y., (2022) Sequential minimal optimization for local scour around bridge piers, *Marine Georesources & Geotechnology*, 40:4, 462-472, DOI: 10.1080/1064119X.2021.1907635

Kim, B.S., & Yoo S. K.,(2006), Motion artifact reduction in photoplethysmography using independent component analysis, *IEEE Transactions on Bio-Medical Engineering*, 53 (3), 566–568.

Kraitl J, Ewald H (2005), Optical non-invasive methods for characterization of the human health status. In: Presented at the 21st international conference on sensing technology, Palmerston North, New Zealand.

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

Lang, N., Kalischek, N., Armston, J., Schindler, K., Dubayah, R., & Wegner, J. D. (2022). Global canopy height regression and uncertainty estimation from GEDI LIDAR waveforms with deep ensembles. *Remote Sensing of Environment*, *268*, 112760.

Lázaro, J., Gil, E., Vergara, J. M., & Laguna, P., (2014), Pulse rate variability analysis for discrimination of sleep-apnea-related decreases in the amplitude fluctuations of pulse photoplethysmographic signal in children. *IEEE Journal of Biomedical and Health Informatics*, 18(1), 240-246.

Liang, Y., Chen, Z., Ward, R., & Elgendi, M. (2018), Photoplethysmography and deep learning: enhancing hypertension risk stratification, *Biosensors*, 8(4), 101.

Liu Z, Zhou B, Li Y, Tang M, Miao F. (2020), Continuous Blood Pressure Estimation From Electrocardiogram and Photoplethysmogram During Arrhythmias. *Front Physiol*. Sep 9;11:575407. doi: 10.3389/fphys.2020.575407. PMID: 33013491; PMCID: PMC7509183.

McDuff D., Gontarek S. & Picard R. W., (2014), Remote detection of photoplethysmographic systolic and diastolic peaks using a digital camera, *IEEE Transactions on Biomedical Engineering*, 61(12), 2948-2954.

O'brien E., Waeber B., Parati G., Staessen J., Myers M.G. (2001), Blood pressure measuring devices: Recommendations of the European Society of Hypertension. *BMJ*.; 322:531–536.

Oğuz, F.E., Alkan, A. & Schöler, T. (2023), Emotion detection from ECG signals with different learning algorithms and automated feature engineering. *SIViP,* 17, 3783–3791. https://doi.org/10.1007/s11760-023-02606-y

Pour Ebrahim, M., Heydari, F., Wu, T. et al. (2019), Blood Pressure Estimation Using On-body Continuous Wave Radar and Photoplethysmogram in Various Posture and Exercise Conditions. *Sci Rep* 9, 16346.

Rodríguez, Juan C.R., et al. (2013), Innovative continuous non-invasive cuffless blood pressure monitoring based on photoplethysmography technology. *Intensive Care Medicine* 39.9: 1618-1625.

Saeed, M., Villarroel, M., Reisner, A.T., Clifford, G., Lehman, L., Moody, G.B., Heldt, T., Kyaw, T.H., Moody, B.E., Mark, R.G.. (2011 May), Multiparameter intelligent monitoring in intensive care II (MIMIC-II): A public-access ICU database. *Critical Care Medicine*, 39(5):952-960; DO: 10.1097/CCM.0b013e31820a92c6.

Salehizadeh, S. M., Dao, D. K., Chong, J. W., McManus, D., Darling, C., Mendelson, Y., & Chon, K. H.,(2014), Photoplethysmograph signal reconstruction based on a novel motion artifact detection-reduction approach. Part II: Motion and noise artifact removal, *Annals of Biomedical Engineering*, 42(11), 2251–2263.

Shin H. S., Lee C., & Lee M., (2009), Adaptive threshold method for the peak detection of photoplethysmographic waveform, *Computers in Biology and Medicine*, 39(12), 1145-1152.

Silva, I, Moody, G. (2014), An Open-source Toolbox for Analysing and Processing PhysioNet Databases in MATLAB and Octave. *Journal of Open Research Software* 2(1):e27 [http://dx.doi.org/10.5334/jors.bi].

Slapničar G, Mlakar N, Luštrek M. (2019), Blood Pressure Estimation from Photoplethysmogram Using a Spectro-Temporal Deep Neural Network. *Sensors* (Basel). Aug 4;19(15):3420. doi: 10.3390/s19153420.

Sun, X., Zhou, L., Chang, S., & Liu, Z. (2021), Using CNN and HHT to predict blood pressure level based on photoplethysmography and its derivatives, *Biosensors*, 11(4), 120.

Sunnetci, KM, Kaba, E, Beyazal Çeliker, F, Alkan, A. (2023), Comparative parotid gland segmentation by using ResNet-18 and MobileNetV2 based DeepLab v3+ architectures from magnetic resonance images. *Concurrency Computat Pract Exper*. ; 35(1):e7405. doi:10.1002/cpe.7405

Suzuki, T., Kameyama, K.,& Tamura, T., (2009), Development of the irregular pulse detection method in daily life using wearable photoplethysmographic sensor, Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Annual International Conference, 6080-6083.

Tanveer, M.S., Hasan, M.K. (2019), Cuffless blood pressure estimation from electrocardiogram and photoplethysmogram using waveform based ANN-LSTM network, *Biomedical Signal Processing and Control*, 51, 382-392.

Tazarv, A. and Levorato, M., (2021), A deep learning approach to predict blood pressure from ppg signals, in 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), pp. 5658–5662, Mexico.

Teng, X. F., and Y. T. Zhang. (2003), Continuous and noninvasive estimation of arterial blood pressure using a photoplethysmographic approach, Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE Cat. No. 03CH37439). Vol. 4. IEEE.

Tjahjadi, H., Ramli, K., & Murfi, H. (2020). Noninvasive classification of blood pressure based on photoplethysmography signals using bidirectional long short-term memory and time-frequency analysis. *IEEE Access*, 8, 20735-20748.

World Health Organization [2022-05-22]. Hypertension. https://www.who.int/news-room/fact-sheets/detail/hypertension

Xing, X., & Sun, M. (2016). Optical blood pressure estimation with photoplethysmography and FFT-based neural networks. *Biomedical Optics Express*, 7(8), 3007-3020.

Yousef Q., Reaz M. B. I., Ali, M. A. M., (2012), The analysis of PPG morphology: investigating the effects of aging on arterial compliance, *Measurement Science Review*, 12(6), 266-271.