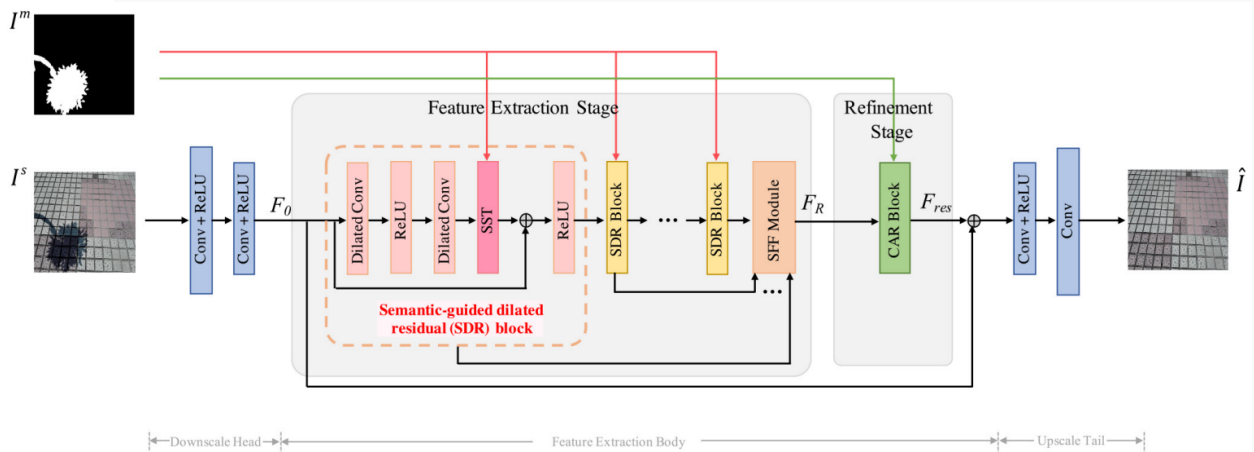**ISSUE: AUGUST 2022**

CTSOC-NCT NEWS ON CONSUMER TECHNOLOGY

ARCHITECTURE OF END-TO-END SHADOW MASK-BASED SEMANTIC-AWARE NETWORK (S2NET) FOR SHADOW REMOVAL

TABLE OF CONTENTS

EDITOR'S NOTE

On behalf of the Editorial Board of IEEE CTSoc News on Consumer Technology (NCT) and my co-editors, Yafei Hou, Luca Romeo, Yuen Peng Loh, Jianlong Fu, I am delighted to introduce the August 2022 issue of the News on Consumer Technology (NCT).

This issue starts with a cover story which presents a computer vision framework, published in IEEE Transactions on Consumer Electronics, as a demonstration of effective shadow removal in images for improving vision tasks such as segmentation and recognition.

Next, an interview with Dr. Sze Ling Tang from Handal Indah Group, Malaysia, presents his experience of industrial adoptions on computer vision, machine learning, and IOT technology. In her point of view, academic research focus on advancing theory to improve efficiency and performance of the work. On the other hand, industry research is more practical work aimed to solve problems, and subsequently improve the quality of life. However, industry research does require the support of academic research to achieve better efficiency and performance in practice.

Besides, I would like to bring to your attention the featured article from Dr. Yung-Hui, Li, founding director of AI Research Center, Hon Hai (Foxconn) Research Institute, presenting his vision on how deep learning revolutionizing mobile automatic blood pressure monitoring.

Happy reading!

Wen-Huang
Cheng
Editor-in-Chief

AUTHOR(S)
Qiqi Bao, Yunmeng Liu, Bowen Gang, Wenming Yang, Qingmin Liao

Shadow regions present a challenging condition for computer vision systems due to their interference on tasks such as segmentation and recognition. Hence, there has been a growing interest in research related to shadow removal. Even so, currently existing powerful deep neural network-based approaches that remove shadows still face difficulties when the shadow regions have similar illumination and color with non-shadow regions, and the presence of artifacts along the boundaries of the shadow. This work proposed a shadow mask-based semantic-aware network (S2net) to address the aforementioned issues. In particular, the proposed end-to-end network can be separated into three parts, a Downscale Head, Feature Extraction Body, and Upscale Tail. The main component proposed is in the Feature Extraction Body that is further split into a Feature Extraction Stage and Refinement Stage. The Feature Extraction Stage consists of a Shadow-mask-based Semantic Transformation (SST) operator that handles the transfer of information to ensure that after shadow removal, the region maintains similar semantics with the non-shadow region features. In the Refinement Stage eliminates the remaining inconsistencies of the shadow removal using guidance from a shadow mask prior. The overall network is then optimized using a proposed Boundary Lost in addition to Pixel, Perceptual, and Color Consistency losses, to eliminate the artifacts along shadow boundaries. The proposed model is able to outperform state-of-the-art algorithms in both quantitative and qualitative measures validating that semantic-awareness and support by prior detected shadow masks can improve shadow removal with potential to boost robustness of computer vision systems.

INTERVIEW WITH DR. SZE LING TANG

Editor: Yuen Peng Loh

Sze Ling Tang received her MSc in Computer Science from Unversiti Sains Malaysia in 2007, and then joined the video analytics team of Mimos Berhad from 2008 to 2011. In the year 2015, she received her Ph.D degree from the Faculty of Computer Science and Information Technology of University of Malaya, under the Department of Artificial Intelligence in 2015. She later join the Handal Indah Group in 2018 and leads the R&D team. Her research areas of interest include computer vision, machine learning and IoT technology.

Could you briefly share about your research and work experiences in this industry?

My research and work experiences have always been focused on computer vision, machine learning and IoT technology. In my current position with a public bus provider company, my role is mainly to lead the team in research and development. In our work, mostly we use our knowledge and experiences to improve the services and products of the company. For instance, we had deployed the "Bus Passenger Counter" to understand passenger flow, the "Visual Vehicle Inspection System" to protect the company's assets, as well as a Recommendation System to devise personalized options for our end users, just to name a few.

What made you join this area of work? Is there a specific motivation or spark that started your interest?

My early interest was in 3D animation when I was young, and so I majored in Computer Aided Geometric Design (CAGD) in university. From there, I wanted to very much apply the varied skill sets I had acquired in the work place but the CAGD market in Malaysia was quite rare during that time. So this lead me to switch my focus on to data visualization and image processing that led into computer vision.

What is the most memorable work that you have done?

My most memorable works have been projects in the industry, related to creating innovative solution to enhance services and improve end user experiences. This is because we are not always required to use or develop the latest and most advanced theory, but instead to creatively think out of the box to figure out practical solution for others. The tip is to truly understand what are the issues faced by end users and to work on their needs.

Based on your experiences, what are the differences between academic research and industry research?

In my point of view, academic research focus on advancing theory to improve efficiency and performance of the work. On the other hand, industry research is more practical work aimed to solve problems, and subsequently improve the quality of life. However, industry research do require the support of academic research to achieve better efficiency and performance in practice.

Was it challenging to transition from academic to industry research and vice versa?

Academic research are typically geared towards proving of advanced concepts where they are usually done in a more controlled setup. Moreover, the measurements used to evaluate the performance are quite standardized. However, in industry research, we need to be prepared that the developed concept or solution is not as straightforward. There may be required to have 20% to 30% of onsite testing or uncontrolled environment fine-tuning process before a developed system can become matured. Also, industry requirement tends to change fast, so working in industry research, in my opinion, would require more problem solving skills as compare to academic research.

How about the roles you have taken? Was it difficult to shift from the role of a researcher into a leader?

Working as a leader in the company, communication has become more important. I need to have close communication with the team and other departments as well because each team or department have their own focus on particular tasks. These tasks are a part of the entire framework or system that we are working on together, so as a leader I am required to manage the progress of the research works to make sure we fulfill the set requirements and within a tight timeline too.

What has been the most eye opening experience you have had in your current role?

In our process of research and development of a solution, we will usually make hypothesis before communicating with users or even conducting a survey to get feedback. I found that the initial solutions we come up with, may not be really preferred by the users especially if there is a lack of understanding on their pain points. My experiences when interacting with the end users to get to know their problems have always been very eye opening to me as we work on understanding their needs.

In your role as a leader, you have met many different people in this line of work. Are there any particular character traits that would excel in this field?

Actually, most who work in the AI industry already have very strong Computer Vision foundation, and starting to excel in Deep Learning algorithms especially object detection. I think this is now the basic technical skills requirement of researchers developing deep learning for industry usage. Recently, my teams are also investigating on active learning by providing more quality data to machine learning in order to improve the performance instead of improving the algorithm or fine-tuning the hyper parameters. What I can see is that they are usually very passionate in what

they are working on, like to understand more and update their knowledge. These are some of the traits that makes someone excel beyond the technical skills.

As the generation gets more tech savvy, do you notice any shift in the abilities or characters of new talents joining the industry?

Yes, the new talents now are more prepared and already equipped themselves for this field with the widely available resources. Some that I have encountered had actively participated in competitions during their studies, some already had their own hands-on experiences with the latest tools, and there are also those that go for tutorials to keep themselves up-to-date with the latest knowledge.

Do you have any words of advice for students and graduates that would be joining this field of work?

We do have trainees from universities join our R&D team. I always like to encourage them to take more initiative. Always think out of the box and voice out your opinions boldly without worrying if the ideas are the best or not. When they show initiative, they can definitely gain more than what they expected.

HOW DEEP LEARNING REVOLUTIONIZING MOBILE AUTOMATIC BLOOD PRESSURE MONITORING

Yung-Hui Li
Founding Director of AI Research Center,
Hon Hai (Foxconn) Research Institute
Yunghui.li@foxconn.com

The Covid-19 pandemic has directly affected thousands of lives. The longer-term effects of COVID-19 are already beginning to emerge: the behavioral health toll of anxiety and depression related to the virus itself, which leads to changing our lifestyle to the new normal. Even though Covid-19 became the number one cause of death in the US several times, we actually faced a long global health burden of cardiovascular diseases (CVDs). CVDs have been a worldwide number one killer that kills more people than Covid-19 and, in ordinary years, more than all other infectious diseases accumulated. In 2600-2700 BCE, during the reign of the Yellow Emperor of China, the wellness of human vital organs was used to be assessed by the pulse. The so-called 'hard pulse disease' has been stated as one of the heart conditions where the pulse hardens due to excessive salt intake in food. People diagnosed with this disease were treated with venesection and bleeding by leeches until 1733 when the term blood pressure (BP) was discovered, and the following pathology of disease related to it embarked on the clinicians' interest [1]. In the modern era, the term hard pulse disease is mainly known as hypertension or high blood pressure. It is a leading preventable risk factor for premature death and disabilities caused by

CVDs. BP dynamics are affected by diets, activities, emotional states, and the use of BP-lowering medication. Changes in some of these factors, such as increasing BMI and stress load, can elevate the BP; in contrast, medication and changes in lifestyle may reduce the raised BP. Though hypertension can be prevented, thus far, people are still doing terribly on BP control worldwide, especially in low-and middle-income countries (see Fig. 1) [2]. Even the awareness of having high BP is less than half the time of the total sufferer. This escalates the number of researchers in developing a comfortable continuous non-invasive BP (CNIBP) measurement system for users.
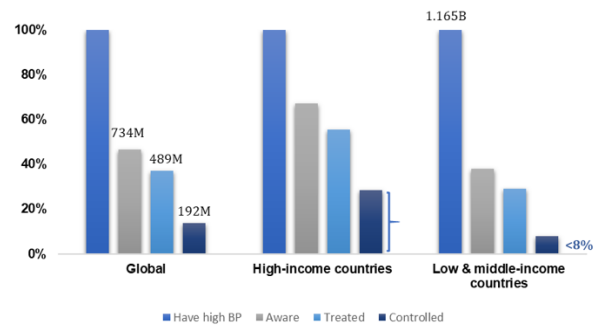

Fig. 1 Most people with hypertension worldwide do not have it under control.

BP Measurement Devices

The invasive catheter system has been considered as the clinical gold standard for measuring continuous BP. This system is performed by physicians or specialized nurses for accurate BP monitoring in intensive care units (ICU). On the other hand, this method is prone to infection when the catheter insertion is performed. Non-invasive cuff-based sphygmomanometer can be an alternative method, which is also recommended for home BP monitoring. Although users can do self-measurement, users have to follow a relatively strict measuring protocol to ensure the values predicted are accurate. The majority (81.4%) of the 914 tested sphygmomanometers exhibited a measuring error that fell within the presently recommended tolerance of ± 3 mmHg. Another drawback is that this device does not allow continuous measurement and the procedure requires time. Both invasive and cuff-based methods are impractical, intermittent, and uncomfortable for patients. Thus, CNIBP systems are expected to fuse the advantages of the two existing methods.

Machine learning-based CNIBP Systems

Several approaches have been proposed, and the pulse transit time (PTT) feature is first utilized. It is often found to have an inverse proportional relationship with BP. By definition, PTT is the travel time between the aortic valve opening and the arrival of the blood flow to the distal location, which can be derived by measuring the time difference between the pulse wave information detected by two sensors apart. There have been a few sensors related to PTT assessment, being investigated in [3]. Based on the heaps of use in literature, the most notable PTT assessment is derived by calculating the time delay between the R peak of the electrocardiogram (ECG) signal to the maximum slope of the (PPG) signal.

At the same time, machine learning regression models, i.e. regression tree, random forest, support vector machine (SVM), help to predict the BP by combining PTT and other related features derived from ECG or PPG signals [4, 5]. Nevertheless, for the overall performance, the prediction error can even be significantly reduced using deep learning methods. Deep learning has a better ability to adapt to represent hierarchical features within multiple layers. We have proven the effectiveness of deep learning techniques compared to some machine learning algorithms in [6]. We proposed a deep long short-term memory (LSTM) model to predict SBP and DBP values from seven features including PTT, heart rate, and the PPG physiology-related information. The real-time demo of this model can be seen in Fig. 2.
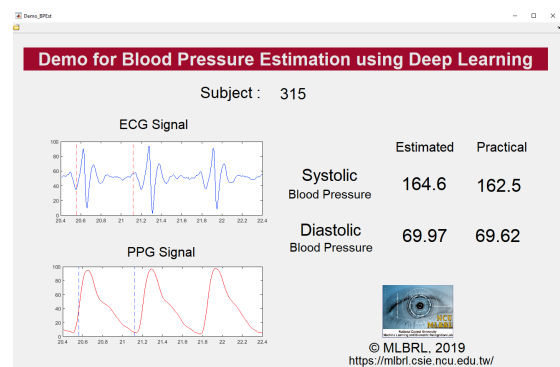

Fig. 1 Our real-time demo program

Deep Learning-based CNIBP Systems using PPG signal only

To accomplish mobile CNIBP systems is a challenging task. Although PPG sensors have been widely used in wearable devices, ECG sensors are still exclusively available in wearable devices. Furthermore, signal retrieval and its preprocessing are the other tricky part. Signals acquired from smartwatches commonly appear quite different from signals acquired from clinical devices due to the very small frequency rate and different kinds of noise that might exist. Thus, a different preprocessing procedure needs to be conducted. To overcome the impracticality of using two separate sensors, most CNIBP system developments are focusing on using

PPG signals only to predict BP. Our initial approach [7] uses a deep neural network (DNN) and 32 selected features from the PPG signal only, illustrated in Fig. 3. Our network architecture contains four hidden layers, denoted as H1, …, H4. The numbers of neurons for H1, H2, H3, and H4 are 2048, 4096, 8192, and 2048, respectively. The first layer contains 32 neurons, corresponding to the number of our features. The last layer includes two neurons for SBP prediction and DBP prediction.
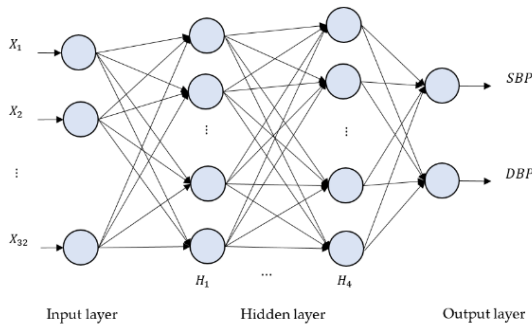

Fig. 2 The deep neural network (DNN) architecture

We decided to adopt the fully connected neural network as our regressor since it is easier to be implemented in wearable devices. The model structure is clean and easier to understand compared to LSTM, which enables software engineers to transfer and deploy the code to wearable devices. Our DNN model achieved a mean absolute error of 3.21 mmHg and 2.23 mmHg for SBP and DBP prediction, respectively.

Commonly, a PPG waveform mainly consists of four distinctive features, namely foot, systolic peak, dicrotic notch, and diastolic peak, as shown in Figure 1. The PPG waveform is quite simple and straightforward but sometimes is not informative. Frequently, the subject's age affects the distinctiveness of the features, such as dicrotic notch, which is usually hard to detect in older subjects, illustrated in Fig. 4. Therefore, features based on dicrotic notch may not be available at all times. Accordingly, we began to focus on developing featureless-based BP estimation.
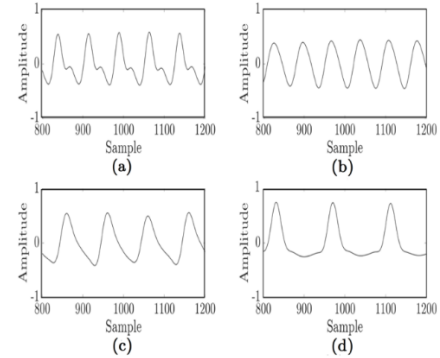

Fig. 4 Waveform variations of PPG waveform including (a) PPG with distinct features, (b) and (c) PPG with indistinct and almost nonexistence dicrotic notch and (d) invisible dicrotic notch with diastolic duration decays faster than the others

Convolutional Neural Network (CNN) is the state-of-the-art of automatic feature extraction while LSTM is an effective choice for analyzing time series data with an ability to handle long sequential data. We proposed a two-hierarchical model consisting of one-dimensional CNN combined with BiLSTM [8]. The lower hierarchy carries out the automatic feature extraction, and the upper learns the temporal relation between the features resulting from the lower part, as illustrated in Fig. 5.
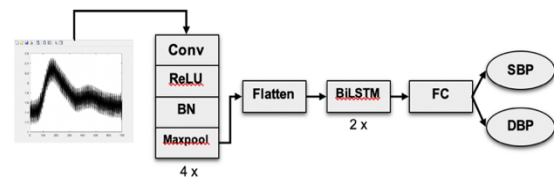

Fig. 5 The 1D CNN-BiLSTM network architecture

Although it did not outperform our DNN model in BP prediction, we believe that in the future, "end-to-end" training, which needs no prior domain knowledge in the loop, will become more popular as the amount of data and computational resources increase. The transition from "feature-based" to "feature-less" signal processing will be a paradigm shift in the biomedical signal processing domain that can also save a lot of training time.

We also proposed a featureless-based model, that not only predicts the SBP and DBP solely but also has the strong learning ability to estimate the whole shape of the arterial blood pressure (ABP)

signal [9]. The input of the proposed model is a raw PPG signal along with its derivatives, instead of the hand-crafted feature of the PPG. This model is unimodal and consists of an LSTM-based autoencoder. Furthermore, we applied transfer learning by first training our autoencoder to reconstruct the PPG waveform input. Then, we freeze the encoding part and only let the next part be trained for constructing the ABP waveform afterward. Taking this application can help our network to learn the intermediate waveform representations explicitly. The training flow of our model is illustrated in Fig. 6
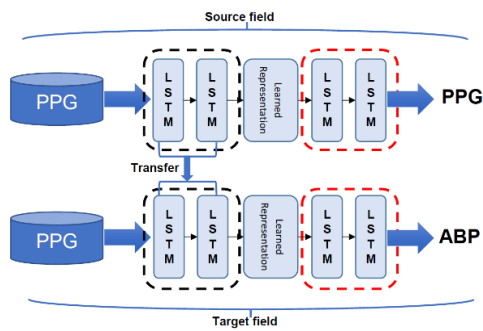

Fig. 6 LSTM-autoencoder training flow. The black dashed-box indicates an encoder, and the red dashed-box indicates a decoder.

The model provides a reasonably accurate and promising result over many subjects being examined, with a mean absolute error of 4.05 mmHg and 2.41 mmHg for SBP and DBP prediction, respectively. Fig. 7 shows the ABP sequence prediction result using the transfer learning method, which has a high resemblance to the observed sequence obtained from the source dataset. In this sense, an LSTM-based autoencoder can perceive the PPG signal information and translate it to the corresponding ABP signal.
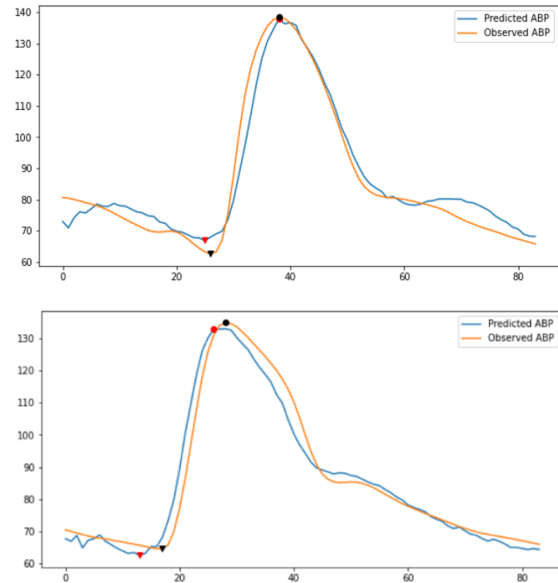

Fig. 3 Examples of ABP prediction results from the proposed model. The circle marks indicate SBP, and the triangle marks indicate DBP.

BPEst Application

Fig. 4 Snapshot of BPEst application for connecting the device with the smartwatch

We have applied our best model to mobile devices as an alternative to the CNIBP system. The application will start working once the smartphone has been connected to a paired smartwatch that has been installed with the same application, as illustrated in Fig. 8. The connection time is less than one minute and the smartwatch will start its PPG

sensor and send the data to the smartphone on the fly. After the smartphone receives the data, the waveform of PPG will be shown on the phone screen, and the prediction process will begin. The prediction result will be shown continuously until the user disconnects the smartwatch or close the app, as shown in Fig. 9. The prediction result will also be displayed on the smartwatch screen with a delay time of less than one second, as shown in Fig. 10. Now, monitoring BP can be performed anywhere and anytime using the technology of deep learning-based mobile BP monitoring system.

Conclusion

Blood pressure control is very important despite being neglected by many people. Monitoring BP regularly can be one effort that could be made. Numerous CINBP system has been developed as deep learning emerges as a robust technology that is extremely beneficial in automatic learning and prediction. We have developed our own deep learning models based on LSTM, DNN, and even CNN. Our best model achieved a similar prediction error with the error tolerance of a sphygmomanometer and applied it to mobile and wearable devices to accomplish a proper mobile BP monitoring system.
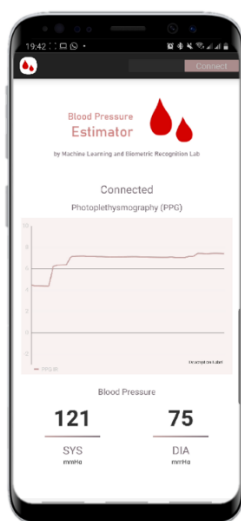

Fig. 6 Snapshot of BPEst application for the smartwatch side

Fig. 5 Snapshot of BPEst application showing the SBP and DBP prediction

References

[1] B. Zhou, P. Perel, G. A. Mensah, and M. Ezzati, "Global epidemiology, health burden and effective interventions for elevated blood pressure and hypertension," Nature Reviews Cardiology, vol. 18, no. 11, pp. 785-802, 2021/11/01 2021, doi: 10.1038/s41569-021-00559-8.

[2] K. T. Mills et al., "Global Disparities of Hypertension Prevalence and Control: A Systematic Analysis of Population-Based Studies From 90 Countries," (in eng), Circulation, vol. 134, no. 6, pp. 441-50, Aug 9 2016, doi: 10.1161/circulationaha.115.018912.

[3] T. Le et al., "Continuous Non-Invasive Blood Pressure Monitoring: A Methodological Review on Measurement Techniques," IEEE Access, vol. 8, 12/04 2020, doi: 10.1109/ACCESS.2020.3040257.

[4] X. Ding, B. P. Yan, Y.-T. Zhang, J. Liu, N. Zhao, and H. K. Tsang, "Pulse Transit Time Based Continuous Cuffless Blood Pressure Estimation: A New Extension and A Comprehensive Evaluation," Scientific Reports, vol. 7, no. 1, p. 11554, 2017/09/14 2017, doi: 10.1038/s41598-017-11507-3.

[5] M. Kachuee, M. M. Kiani, H. Mohammadzade, and M. Shabany, "Cuff-less high-accuracy calibration-free blood pressure estimation using pulse transit time," in 2015 IEEE International Symposium on Circuits and Systems (ISCAS), 24-27 May 2015 2015, pp. 1006-1009, doi: 10.1109/ISCAS.2015.7168806.

[6] Y.-H. Li, L. N. Harfiya, K. Purwandari, and Y.-D. Lin, "Real-Time Cuffless Continuous Blood Pressure Estimation Using Deep Learning Model," Sensors, vol. 20, no. 19, p. 5606, 2020. [Online]. Available: https://www.mdpi.com/1424-8220/20/19/5606.

[7] Y.-C. Hsu, Y.-H. Li, C.-C. Chang, and L. N. Harfiya, "Generalized Deep Neural Network Model for Cuffless Blood Pressure Estimation with Photoplethysmogram Signal Only," Sensors, vol. 20, no. 19, p. 5668, 2020. [Online]. Available: https://www.mdpi.com/1424-8220/20/19/5668.

[8] Y.-H. Li, L. N. Harfiya, and C.-C. Chang, "Featureless Blood Pressure Estimation Based on Photoplethysmography Signal Using CNN and BiLSTM for IoT Devices," Wireless Communications and Mobile Computing, vol. 2021, p. 9085100, 2021/11/26 2021, doi: 10.1155/2021/9085100.

[9] L. N. Harfiya, C.-C. Chang, and Y.-H. Li, "Continuous Blood Pressure Estimation Using Exclusively Photopletysmography by LSTM-Based Signal-to-Signal Translation," Sensors, vol. 21, no. 9, p. 2952, 2021. [Online]. Available: https://www.mdpi.com/1424-8220/21/9/2952.