Deep learning algorithm evaluation of hypertension classification in less photoplethysmography signals conditions

Chih-Ta Yen¹, Sheng-Nan Chang² and Cheng-Hong Liao³

Abstract
This study used photoplethysmography signals to classify hypertensive into no hypertension, prehypertension, stage I hypertension, and stage II hypertension. There are four deep learning models are compared in the study. The difficulties in the study are how to find the optimal parameters such as kernel, kernel size, and layers in less photoplethysmography (PPG) training data condition. PPG signals were used to train deep residual network convolutional neural network (ResNetCNN) and bidirectional long short-term memory (BiLSTM) to determine the optimal operating parameters when each dataset consisted of 2100 data points. During the experiment, the proportion of training and testing datasets was 8:2. The model demonstrated an optimal classification accuracy of 76% when the testing dataset was used.

Keywords
Photoplethysmography, hypertensive, deep learning, residual network convolutional neural network, bidirectional long short-term memory

Introduction
Signals acquired using photoplethysmography (PPG) are referred to as PPG signals. By using light sensors to absorb light energy, PPG signals record signals generated by blood flow variations in blood vessels. Intravascular blood flow (per unit) changes according to heart pulse, and the inductive voltage detected by light sensors changes according to blood flow variations. Light sensors absorb the most light during systole. Therefore, PPG signal amplitudes are proportional to variations of blood influxes and effluxes in tissue. PPG signals are easy to acquire; after undergoing proper signal preprocessing, they have high use value. A study indicated that PPG and electrocardiography (ECG) are related.¹ Additionally, ECG and PPG signal collection and analysis systems have been developed to assess and analyze the correlation between PPG and ECG. Numerous studies have combined PPG and ECG signals in research and proven the applicability of ECG and PPG analysis in investigating arterial changes caused by age.² In Liou and Su,³ a system was proposed that extracts attributes from ECG and PPG signals to estimate the systolic and diastolic blood pressure (BP) of each heart pulse. Furthermore, the research Şentürk et al.⁴ proposed a novel algorithm that removes noise in highly unstable ECG and PPG signals to acquire pulse rate and measure cardiovascular parameters in a timely manner (e.g. pulse rate, total cardiac cycle, and BP). These parameters are applicable for the continuous monitoring of health conditions in patients with cardiac diseases. In Yude et al.,⁵ according to two variables' AR model were employed in a developed algorithm to verify that the PPG spectrum consists of the respiration rate; this indicated the possibility of using PPG signals to acquire respiration and heart rates. In research Şentürk et al.,⁴ ECG and PPG signals were measured and a repetitive neural network was employed to propose a novel hybrid prediction model to continuously predict BP. In

¹Department of Electrical Engineering, National Taiwan Ocean University, Keelung City
²Division of Cardiology, Department of Internal Medicine, National Taiwan University, Yun-Lin Branch, Dou-Liu City
³Department of Electrical Engineering, National Formosa University, Yunlin County

Corresponding author:
Chih-Ta Yen, Department of Electrical Engineering, National Taiwan Ocean University, No. 2, Beining Rd., Zhongzheng Dist., Keelung City 202301, Taiwan (R.O.C.).
Email: chihtayen@gmail.com

Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (https://creativecommons.org/licenses/by/4.0/) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).
Dutt and Shruthi, two different methods were employed to acquire arterial parameters; in one method, ECG and PPG signals were simultaneously acquired and the synchronized signals were used to obtain the pulse wave transmission time, which was subsequently utilized to measure the signal strength. In the other method, PPG signals were used to obtain arterial parameters (e.g. the stiffness index and augmentation index). Subsequently, a spectral power analysis was used to indicate differences in frequencies between two age groups. This research proved the usability of ECG and PPG analysis in investigating arterial characteristic changes caused by age and its potential applicability in the field of health care. In Verma and Agarwal, an ear-worn monitor was developed to measure long-term BP and heart rate (HR). This proof-of-concept system demonstrated promising feasibility in the use of ear-ECG/PPG to measure exercise tolerance through BP/HR monitoring. In addition, researchers have used PPG signals to understand sympathetic nervous system changes and support cardiovascular function diagnosis. When signal sources are stable, PPG signals have outstanding potential to be used to monitor HR and HR variability (HRV).

Photoplethysmographs are the most common wearable device for measuring physiological signals. However, current studies on PPG have focused on BP estimation accuracy, sleep quality analysis, and emotional and personality judgment. For example, Shimazaki et al. used PPG physiological signals to acquire the blood vessel stiffness index and used it to diagnose arterial stiffness severity. In Dutt and Shruthi, an artificial neural network was employed to develop a novel method that used BP to estimate PPG signals. Regression calculations showed that the estimation accuracy of systolic and diastolic BP was 98.666% for both. Studies on PPG are widespread; Şentürk et al. employed PPG to replace the ECG method in HRV analysis. This enabled the researchers to analyze the sympathetic nerve and parasympathetic nerve activities in automatic nerve systems, thereby allowing the objective assessment of mental fatigue. In Manimegalai et al., convolutional neural network technology was proposed to estimate BP using the formation characteristics of pulse waves. However, compared with public databases for disease classification data, databases that provide PPG measurement data are sparse; researchers frequently encounter the challenges of insufficient data length and the need to organize data. Therefore, this study used limited available data to develop a deep learning technology that assesses network architecture and network parameter allocation to enhance overall classification accuracy when data available are limited. The less data condition is not suitable for deep learning algorithm, how to find the optimal parameters of the models and the precision data format might be the important issues in the future. The other research with deep learning method is also presented, for example an application-based online and offline traffic classification. The remainder of this paper is organized as follows. Section 2 describes the related work of PPG signals. Section 3 presents the deep learning models, which involved LSTM, Xception, ResNetCNN + BILSTM and Xception + BILSTM. Section 4 presents and discusses the results of the several deep learning models. Section 5 concludes this paper.

Related work

Figure 1 depicts a PPG signal, which comprises the U point, P wave, T wave, V wave, and D wave.

1. U point (UP stroke): This wave reflects the end-diastolic intravascular pressure and volume, thereby serving as the initial point when the ventricle rapidly ejects blood, which flows into the artery. This point is also called the “flowing point.”

2. P wave (percussion wave): This wave is the main amplitude of the pulse signal figure caused by the substantial ventricle blood output during systole. The wave rising speed is affected by cardiac output, ventricular ejection velocity, arterial resistance, and the elasticity of the vessel wall.

3. T wave (tidal wave): This wave reflects when the left ventricle stops ejecting blood, the artery expands and reduces pressure, and an antidiromic reflected wave of blood flows in the artery. This wave is influenced by the degree of variety of peripheral vascular resistance, elasticity of the vessel, and descent velocity of the recession limb.

4. V wave (valley wave): This wave represents the emptying time of the aortal static pressure and serves as the cut-off point of systole and diastole. This wave is affected by peripheral vascular resistance and the descent velocity of the recession limb.

5. D wave (dicrotic wave): This wave represents when the aortic valves close suddenly in the early diastole phase of the ventricular. Then,
blood that flows back affects it before being bounced back, which causes aortal pressure to rise again. This wave is primarily influenced by the elasticity of the vessel and the state of the closed aortic valves.

The aforementioned five components of PPG signals can be used to distinguish PPG characteristics. Therefore, researchers have focused on inspecting these characteristics and employed classification algorithms to acquire effective benefits in PPG characteristics. Because PPG is a time series, signal processing technology can be used to extract necessary characteristics from the signals. For example, discrete wavelet transform, discrete Fourier transform, or continuous wavelet transform can be used to extract meaningful characteristics. Ferdinando et al.\textsuperscript{13} detailed meaningful characteristics that can be acquired through PPG and signal processing technology. Once the extraction process was complete, the extracted characteristics could be used to produce training data and applied in machine training algorithms (e.g. support vector machine and \textit{k}-nearest neighbor) to construct prediction models.

To utilize signals in machine learning, the signal characteristics must first be extracted. Extracted characteristics that are more related to the goal exhibit greater classification accuracy. However, the main disadvantage of this method is that the characteristics used for machine learning must be extracted manually. Therefore, researchers must first research which characteristics have a high correlation with the goal.

In recent years, deep neural networks have garnered interest in the field of biomedical signal processing application. Convolutional neural networks (CNNs), capable of directly applying raw ECG signals, do not require the signals to undergo filtering technology or characteristic extraction beforehand; they directly extract transient characteristics from ECG signals to complete the classification of cardiovascular diseases. In Murugesan et al.,\textsuperscript{14} a CNN was employed to integrate single-lead ECG signals with long short-term memory (LSTM) algorithms to classify arrhythmia symptoms (e.g. normal beat, premature ventricular contraction, and premature atrial contraction), achieving a classification accuracy of 97.6%. In Chen et al.,\textsuperscript{15} a CNN was used to integrate 12-lead ECG signals with bidirectional LSTM (BILSTM) algorithms for the classification of atrial fibrillation, atrial flutter, complete atrioventricular block, junctional rhythm, sinus node diseases, and Wolff–Parkinson–White syndrome, achieving an overall mean accuracy of 81%. Finally, Peng et al.\textsuperscript{16} employed a CNN and LSTM to enhance classification accuracy and extract characteristics from denoised and filter-processed ECG signals to classify arrhythmia symptoms. The model had a classification accuracy, sensitivity, and specificity of 99.12, 99.15, and 99.84%, respectively, in the classification of the five types of ECG signals (i.e. nonectopic, supraventricular ectopic, ventricular ectopic, fusion, and unknown).

**Proposed deep learning method**

This study used PPG time series data as the input data for its experiment. In time series data, correlation between permutation sequences is crucial to model performance. Therefore, the time information in time series data codes and the wave characteristics in the measured signal space must be simultaneously extracted to enhance classification performance. This study evaluated the performance of two types of deep learning models, namely ResNetCNN + BILSTM (Figure 2) and Xception + BILSTM (Figure 3), in classifying cardiovascular diseases according to the input data. According to database information, each input data consisted of 2,100 data points. According to different measurers, each of these 2,100 data points could be used to draw out 2 to 3 PPG waves. We first use ResNetCNN + BILSTM models, which has shown the good performance of PPG signals in recent researches. The difficulties in the study are how to find the optimal parameters such as kernel, kernel size and layers in less PPG training data condition. And the second, we propose our model of Xception + BILSTM and also find the optimal parameters to keep the accuracy in the same PPG database.

In the ResNetCNN + BILSTM model, a 36-layer ResNetCNN was employed to extract characteristic vectors from the input data. After extraction, each data point yielded a characteristic vector of the size $15 \times 132$. Subsequently, the characteristic vectors were input into BILSTM for the machine to learn the time relation between the extracted characteristics. In the Xception + BILSTM model, the 37-layered Xception was used to extract characteristics from the input data. After extraction, each data point yielded a characteristic vector of the size $64 \times 132$. Subsequently, the characteristic values were input into BILSTM for the machine to learn the time relationship between the
characteristics. The goal of this study was to classify hypertensive disease test results using PPG data. Therefore, cross entropy loss was selected as the loss function of the networks.

### Experiments of the proposed deep learning models

**Dataset**

This study employed the database established by Liang et al., which consists of data collected from 219 participants with ages in the range of 20–89 years from Guilin People’s Hospital in mainland China. We concern the less dataset condition of deep learning models. Although there are several schemes, for example, pre-trained model and generative model to overcome less dataset condition, but it still doesn’t have confidence in practical applications. The database provided 657 sets of short-length PPG data segments and included information on the participants’ sex, age, weight, height, BMI, systolic pressure, and diastolic pressure. PPG signals were measured at a 1 kHz sampling frequency and each data segment was 2.1 s in length. The data consisted of four types of hypertension diagnosis: healthy, prehypertension, stage I hypertension, and stage II hypertension. The abstract of the data set is listed in Table 1. The following section describes how different combinations of network parameters were employed to evaluate the optimal performance of the network system.

![Figure 3. Proposed deep learning architecture for Xception + BILSTM.](image)

**Experimental setting**

The experiment software included Anaconda, Spyder, and Keras, which were built on top of TensorFlow. The study focus on how to choose the optimal model and parameters in less training data condition. In the deep learning models, we tune three kinds of parameters, for example, layers, kernels, and kernel size in different deep learning models. We believe that with less training continuous PPG signals, the performance will depend on spatial and time features. Hence, we combine ResNetCNN or Xception models with BILSTM model to increase accuracy. In addition, this study employed the ResNetCNN, Xception, and BILSTM network architectures as deep learning models. Different parameters were input into ResNetCNN and Xception, and the networks were used to extract data characteristics. Subsequently, one-layer BILSTM was used to extract time transmission characteristics from the PPG wave.

The simulated parameters are presented in Tables 2 and 3. We not only focus on how to find the optimal parameter combinations in different models, but also consider spread and deep models to analyze which is more suitable for one dimensional data of deep learning algorithm. The tables list the different parameters of ResNetCNN and Xception, respectively. The criterions in less data training scenario is first to find the appropriate number of layers and then searching kernel and its kernel size. In addition, to assess the performance of the neural networks when operating with less data, this study compared the classification performance of ResNetCNN when operating under different parameters, namely layers = 21, 26, 31, 36, and 41; when the kernel = 13, 14, 15, 16, 17 at layers = 31; and when kernel size = 37, 38, 39, 40, 41 when kernel = 15 and layers = 31. Through this, this study aimed to acquire the optimal parameter for the network to conduct data characteristic extraction. Similarly, this study evaluated the classification performance of Xception in

| Class                | Number of samples |
|---------------------|-------------------|
| Normal              | 240               |
| Prehypertension     | 253               |
| Stage 1 hypertension| 120               |
| Stage 2 hypertension| 61                |
| Total               | 654               |

| ResNetCNN            |
|----------------------|
| Layers | 21 | 26 | 31 | 36 | 41 |
| Kernel | 13 | 14 | 15 | 16 | 17 |
| Kernel size | 37 | 38 | 39 | 40 | 41 |

The abstract of the data set is listed in Table 1. The following section describes how different combinations of network parameters were employed to evaluate the optimal performance of the network system.

| Table 1. Summary of the dataset. |
|-----------------------------------|
| Class                | Number of samples |
|----------------------|-------------------|
| Normal               | 240               |
| Prehypertension      | 253               |
| Stage 1 hypertension | 120               |
| Stage 2 hypertension | 61                |
| Total                | 654               |

| Table 2. Different ResNetCNN parameters. |
|-----------------------------------------|
| ResNetCNN                               |
| Layers | 21 | 26 | 31 | 36 | 41 |
| Kernel | 13 | 14 | 15 | 16 | 17 |
| Kernel size | 37 | 38 | 39 | 40 | 41 |
different parameter designs, the first kind is layers = 31, 34, 37, 40, and 43; when the kernel = 30, 31, 32, 33, 34 at layers = 37; and when kernel size = 34, 35, 36, 37, 38 when kernel = 32 and layers = 37.

Finally, this study evaluated the accuracy, recall, and precision of the models when operating under different parameters.

**Experimental results**

Training and testing datasets were randomly distributed at a ratio of 8:2. In the experiment, the network performance efficiencies of the deep learning models (i.e. LSTM, ResNetCNN, Xception, ResNetCNN + BILSTM, and Xception + BILSTM) were compared, and the importance of time series and network spatial characteristics in the PPG data was evaluated. After verifying the results, this study employed different parameters to analyze the network performance of ResNetCNN + BILSTM and Xception + BILSTM.

The experiment results revealed that after 200 iterations, the characteristics extracted using ResNetCNN + BILSTM and Xception + BILSTM were more accurate compared with those extracted by ResNetCNN, Xception, or LSTM alone. The Table 4 shows the results in different models. When ResNetCNN or Xception was used alone, the measured datasets were uniformly classified as prehypertension. By contrast, the LSTM network architecture alone demonstrated unfavorable resolution performance. The Table 4 shows among the different single models of Xception, ResNetCNN and BILSTM. Although the accuracy of only ResNetCNN is up to 73%, but the Recall is only 10% and the precision is also too low.

The Table 5 shows among the different parameter combinations input into ResNetCNN + BILSTM, when the layers could be 21, 26, 31, 36, and 41, the model produced the most satisfactory results at layer = 31. The reason is the accuracy of layers = 21, 26, and 31 are different from 1%, but the precision of layer = 31 is higher 7% than layers = 21 and 26. When the layer = 31 and the kernel could be 13, 14, 15, 16, and 17, the model demonstrated optimal results at kernel = 15. This is because of the precision of kernel = 13 is higher than kernel = 15, but the recall and accuracy are both lower than kernel = 15. Finally, when the layer = 31 and kernel = 15, the kernel size could be 37, 38, 39,
40, and 41 the model had optimal results at kernel size = 38. The criterion is that we can keep the optimal parameters between all accuracy, recall and precision.

The same logic of Table 6 shows among the different parameter combinations input into Xception + BILSTM, when the layers could be 31, 34, 37, 40, and 43, the model produced the most satisfactory results at layer = 34. When the layer = 37 and the kernel could be 30, 31, 32, 33, and 34, the model demonstrated optimal results at kernel = 32. Finally, when the layer = 37 and kernel = 32, the kernel size could be 34, 35, 36, 37, and 38 the model had optimal results at kernel size = 36.

Figures 4 and 5 demonstrate that, after 200-epoch training, the accuracy of the validation set and of the model both become more stable. The simulation parameters are the same as Figure 4. Much noise was observed, because the data of this study had not undergone filtering, leading to inferior results from the open dataset. However, in terms of overall model accuracy, the Xception + BILSTM this study proposed was effectively verified.

Figure 6 shows that accurate classification is achieved in four types of hypertension diagnosis: healthy, prehypertension, stage I hypertension, and stage II hypertension. The parameters of Xception + BILSTM model are kernel = 32, kernel size = 36, and Xception layers = 37. In general, data classification issue of deep learning model is always limited in data unbalanced conditions. How to find the optimal parameter combinations become an important key point in the study? The confusion matrix shown seven data entries for stage I hypertension were classified as prehypertension and two data entries for stage II hypertension were classified as prehypertension. The confusion matrix indicates that stage I hypertension and stage II hypertension in the static actions were not wrongly classified. The reason for this is that significant changes in the numerical value of the PPG measurement occur during stage I hypertension and stage II hypertension were susceptible to wrong classification in the model because the respective data were more similar, and the training data were less in the open dataset.

**Conclusion**

This study employed different deep learning models to analyze how the classification accuracy rate of hypertensive disease test results can be enhanced through the selection of different models and parameters given limited available PPG signal data. Research analysis and comparison revealed that when the data consisted of only 2100 data points, the ResNetCNN + BILSTM model had optimal results when the layers = 31, kernel = 15 and the kernel size = 38. In particular, the model had an accuracy of 74%, 39% recall, and a
precision of 49%. By comparison, the Xception + BILSTM model had optimal performance when the layers = 37, kernel = 32, and kernel size = 36, resulting in an accuracy of 76%, 45% recall, and a precision of 48%. This indicated that the results of the Xception + BILSTM model were more satisfactory compared with those of the ResNetCNN + BILSTM model in less data training condition. However, because public PPG databases are sparse, the data quantity was insufficient. Therefore, this study could only compare the of hypertensive disease classification accuracy of two deep learning models. In the future, the researchers hope to enhance the classification accuracy rate of the deep learning models through increasing testing data or employing suitable data expansion techniques.

Declaration of conflicting interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported in part by the Ministry of Science and Technology MOST 108-2221-E-150 -022 -MY3.

ORCID iD
Chih-Ta Yen https://orcid.org/0000-0003-0831-6774

References
1. Hsiao C, Lee R-G, Hsu F-W, et al. Correlation analysis of heart rate variability between PPG and ECG for wearable devices in different postures. In: 2017 IEEE international conference on systems, man, and cybernetics (SMC), Banff, AB, 2017, pp.2957–2962. doi: 10.1109/SMC.2017.8123077.
2. Hernando A, Pelaez-Coca MD, Lozano MT, et al. Autonomic nervous system measurement in hyperbaric environments using ECG and PPG signals. IEEE J Biomed Health Inform 2019; 23(1): 132–142.
3. Liu J and Su T-Y. An investigation of the ECG and PPG signal acquisition analysis system. In: 2018 IEEE international conference on consumer electronics-Taiwan (ICCE-TW), Taichung Taiwan, 2018, pp. 1–2, doi: 10.1109/ICCE-China.2018.8448500.
4. Şentürk Ü, Vicedağ I and Polat K. Repetitive neural network (RNN) based blood pressure estimation using PPG and ECG signals. In: 2018 2nd International symposium on multidisciplinary studies and innovative technologies (ISMSIT), Ankara Turkey, 2018, pp. 1–4, doi: 10.1109/ISMSIT.2018.8567071.
5. Yude L, Liu W-T, Tsai C-C, et al. Coherence analysis between respiration and PPG signal by bivariate AR model. Proc World Congr Sci Eng Technol 2009; 41: 847–852.
6. Dutt DN and Shruthi S. Digital processing of ECG and PPG signals for study of arterial parameters for cardiovascular risk assessment. In: 2015 International conference on communications and signal processing (ICCSP), Melmaruvathur, India, 2015, pp. 1506–1510. IEEE doi: 10.1109/ICCSP.2015.7322766.
7. Verma D and Agarwal S. Cardiac arrhythmia detection from single-lead ECG using CNN and LSTM assisted by oversampling. In: 2018 International conference on advances in computing, communications and informatics (ICACCI), Bangalore India, 2018, pp.14–17, 2018. IEEE. doi: 10.1109/ICACCI.2018.8554541.
8. Shimazaki S, Kawanaka H, Ishikawa H, et al. Cuffless blood pressure estimation from only the waveform of photoplethysmography using CNN. Annu Int Conf IEEE Eng Med Biol Soc 2019; 2019: 5042–5045.
9. Şentürk Ü, et al. Cuffless continuous blood pressure estimation from Electrocardiogram (ECG) Photoplethysmography (PPG) signals with artificial neural network. In: 2018 26th Signal processing communications applications conference (SIU), Izmir Turkey, pp.1–4, 2018. doi: 10.1109/SIU.2018.8404255.
10. Manimegalai P, Jacob D and Thanushkodi K. Estimation of cardiovascular parameters from ECG and PPG signals. In: 2012 38th Annual northeast bioengineering conference (NEBEC), Philadelphia, PA, pp.237–238, 2012, doi: 10.1109/NEBEC.2012.6207051.
11. Lee D, Lee S-J and Seo Y-J. Application of recent developments in deep learning to ANN-based automatic berthing systems. Int J Eng Technol Innov 2020; 10(1): 75–90.
12. Chang L-H, Lee T-H, Chu H-C, et al. Application-based online traffic classification with deep learning models on SDN networks. Adv Technol Innov 2020; 5(4): 216–229.
13. Ferdinando H, Huotari M and Myllyla T. Photoplethysmography signal analysis to assess obesity, age group and hypertension. Annu Int Conf IEEE Eng Med Biol Soc 2019; 2019: 5572–5575.
14. Murugesan B, Ravichandran V, Ram K, et al. ECGNet: deep network for arrhythmia classification. In: 2018 IEEE international symposium on medical measurements and applications (MeMeA), Rome Italy, pp.1–6, 2018. IEEE. doi: 10.1109/MeMeA.2018.8438739.
15. Chen Y, Liu C, Tseng VS, et al. Large-scale classification of 12-lead ECG with deep learning. In: 2019 IEEE EMBS international conference on biomedical & health informatics (BHI), Chicago, IL, USA, 2019, pp.1–4. IEEE. doi: 10.1109/BHI.2019.8834468.
16. Peng X, Shu W and Song WW. Automatic classification of ECG signals in WBAN based on convolutional neural network and long–short term memory network. In: 2019 4th International conference on computational intelligence and applications (ICCIA), Nanchang, China, pp.104–111, 2019. doi: 10.1109/ICCIA.2019.00027.
17. Liang Y, Chen Z, Liu G, et al. A new, short-recorded photoplethysmogram dataset for blood pressure monitoring in China. Sci Data 2019; 5(180020): 1–7.