Deep Learning Based Heart Rate Estimation for Photoplethysmogram Signals

Seong-Hyun Kim [1] and Eui-Rim Jeong [1+]
[1] Graduate Student, Dept. of Mobile Convergence and Engineering, Hanbat National University, Daejeon, Republic of Korea
[1+]Professor, Dept. of Information and Communication Engineering, Hanbat National University, Daejeon, Republic of Korea

Article Info
Page Number: 479 – 486
Publication Issue:
Vol. 71 No. 3 (2022)

Abstract
Heart Rate (HR) is a basic and importantbiomarker that measures heart beat rates. This paper proposes HR estimation technique based on deep learning technique using Photoplethysmogram (PPG) signals. For reliable and accurate HR estimation, we propose a 1-Dimensional Convolutional Neural Network (1D-CNN) model which consists of 10 convolutional layers and 2affine layers. To examine the HR estimation accuracy, cross validation for all possible combination of training and test data sets is performed. The programming tools are Python 3.7.5 andKeras 2.0. To avoid overfitting due to the small data set, data augmentation technique is incorporated. The loss function for training is the Mean Square Error (MSE), one of the famous errors for a regression problem. For verification performance, we measure Mean Absolute Error (MAE). According to the verification results, the proposed estimator shows an MAE of 1.23 Beats Per Minute (BPM). This results indicate that the proposed technique can be used as an alternative of the existing techniques. If the technique is applied to wearable devices, reliable day-and-night HR measuring can be possible.

Article History
Article Received: *12 January 2022*
Revised: *25 February 2022*
Accepted: *20 April 2022*
Publication: *09 June 2022*

Keywords:1-dimensional, CNN, deep learning, heart rate, PPG

1. Introduction

Heart Rate (HR) is an important biomarker for representing health conditions and the unit is Beats Per Minute (BPM) (Acharya et al., 2006). For a normal person, the HR is around 70 to 80 BPM in calm condition (Christofaro et al., 2017). HR is highly dependent of activities such as sleep, breathing, and physical exercise (Reimers et al., 2018). HR also depends on diseases (Kim et al., 2018). For instance, for people with arrhythmia, HR can increase abruptly (Hannun et al., 2019). High HR People have a higher possibility of cardiovascular disease and accordingly, the mortality rate can be higher than low HR people (Grande et al., 206). Therefore, constant monitoring HR and finding HR anomalies quickly is important.

+ Corresponding author. Tel.: +82-42-821-1752; fax: +82-42-821-1595.
E-mail address: erjeong@hanbat.ac.kr.

HR can be measured fromElectrocardiogram (ECG) or Photoplethysmogram (PPG) signals (Lu et al., 2009). Between them, PPG signal is easier to get with simple sensors. The PPG sensor measures the intensity of the reflected light after transmitting the light emitting diode (LED) to the skin. Information on the change in blood flow can be found in PPG signals (Temko, 2017). One beat of PPG has two peaks: systolic peak and diastolic peak (Yousef et al., 2012). Systolic peak is the peak corresponding to the heart pumping and is due tothe direct pressure wave moving from the left ventricle to the end of the body. The second peak is the result of reflection of the first wave by arteries. Therefore, the diastolic peak is much smaller than the systolic peak. Rarely, some people may have very large diastolic peak. HR can be measured by counting the systolic peaks, but when the diastolic peak is large and comparable to the systolic peak, double HR can be measured by counting both the systolic and diastolic peaks. Thus, care should be taken not to measure the diastolic pulse.

This paper proposes a 1-Dimension Convolutional Neural Network (1D-CNN) (Eren et al., 2018)(Kwon et al., 2020)that estimates HR reliably from the raw PPG signals. Deep learning techniques such as 1D-CNN learns from data, and if a lot of and various data are prepared, the deep learning can predict the HR reliably and stably. The PPG signals used in this paper have both normal signals with small diastolic peaks and abnormal signals with large diastolic peaks. Since the number of abnormal data can be insufficient, data augmentation is employed to increase the number of data. The model of the proposed 1D-CNN is composed often convolutional layers and two affine layers. To avoid any performance bias due to data selection, cross validation is performed for all possible data combination of training and test (Browne, 2000). According to the final performance verification through cross validation, the proposed technique hasan MAE of 1.23 BPM which is small enough to be used as a HR monitor. Thus, if the proposed technique is applied to wearable devices such as smart band or watch, HR can be accurately measured day-and-night.

2. System Model

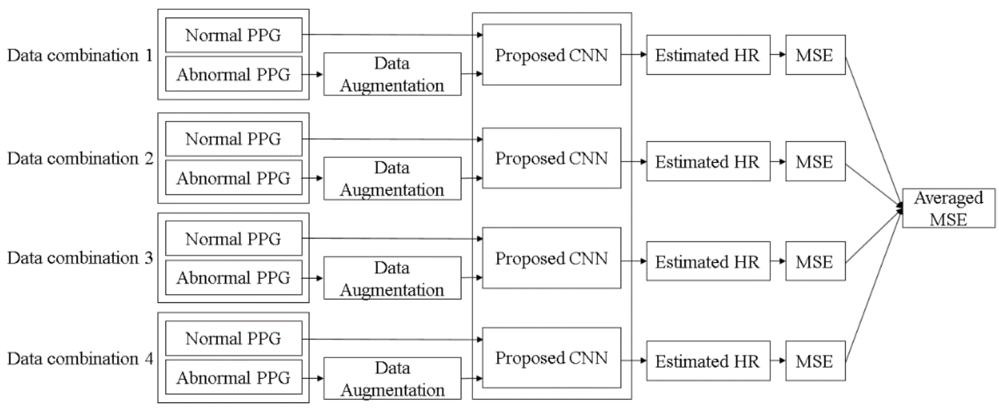


Fig.1: Block Diagram of Proposed HR Estimation Technique

Fig. 1 shows the system block diagram for HR estimator proposed in this paper. First, PPG signalsare divided into training and validation, and four combinations are used for cross validation. Among PPG signals, those with abnormally large diastolic peaks are fewer than the normal signals so that the number of abnormal PPG signals increases by data augmentation technique. For all the augmented abnormal PPGs and normal PPGs, the proposed 1D-CNN is trained. The cross validation avoids the performance bias across the combinations of the training and test data. For all possible combinations, the verification performance is averaged. The averaged performance can demonstratemore reliable result.

3. PPG Signal

In this paper, HR estimation is based on the PPG signals. PPG is one of the methods of measuring movement of blood vessels. PPG measures the signal between the light source and the light receiver. Specifically, the PPG sensor measures the intensity of the reflected light after transmitting the light emitting diode (LED) to the skin. Information on the change in blood flow can be found in PPG signals.Therefore, it is possible to know heartbeat from the amount of blood flow changes. This is the reason how the HR measuring is possible with the PPG signals. The data used for training and verification are total 28,378 normal PPGs and 188 abnormal PPGs. Each PPG signal is sampled with 50 Hz clock frequency and 10 seconds long.Consequently, the length of each PPG signal is 50 x 10 = 500 samples.

3.1. Diastolic peak

The largest peak of PPG signal is a systolic peak. Systolic peak is the output of a direct heart pressure moving from the left ventricle to the end of the body. In addition to the systolic peak, there is a small diastolic peak as well. Diastolic peak is the result of reflection of the systolic wave by arteries in the end of body. Fig. 2 shows examples for a normal PPG (left) and an abnormal PPG (right) where diastolic peaks are large, respectively.

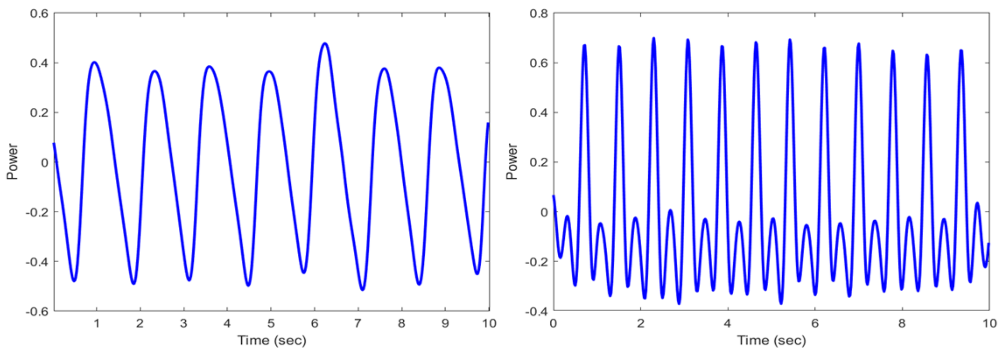


Fig. 2: Two case of PPG: normal PPG (left) and abnormal PPG (right)

When the diastolic peak is unusually large as shown in the PPG on the right side of Fig. 2, not only the systolic peak but also the diastolic peak can be counted when HR calculation. In this case, a problem may happen that the HR estimate is twice the actual HR. Therefore, to identify the abnormal diastolic peak and estimate HR accurately, a lot of PPG signals for those abnormal signals should be prepared. To this end, data augmentation is performed to increase the number of abnormal PPG signals.

4. Data Augmentation

Data augmentation is performed to prepare the training and verification data for the proposed 1D-CNN. The objective of data augmentation is to increase the number of the abnormal diastolic PPG signals to make the numbers of normal PPGs and abnormal PPGs comparable. Data augmentation in image identification fields is a technique that increases the number of images by takingsmall changes such as rotation, zoom-in, zoom-out, and mirror symmetry. Data augmentation is anessential tool when there are only a few data in hand. For data augmentation, a window function is introduced. The window function used for data augmentation is written in (1).

$$W(n) = P * (sin(2 * \pi * F_c * n + Offset) + 10) \text{ for } n = 0, \ldots, 499 \tag{1}$$

In (1), P is randomly selectedfrom 0.06 to 0.15, F_c is from 0.3 to 3, and the Offset is from 0.01 to 1. Some example shapes of $W(n)$ that is madeby (1) are shown in Fig. 3.

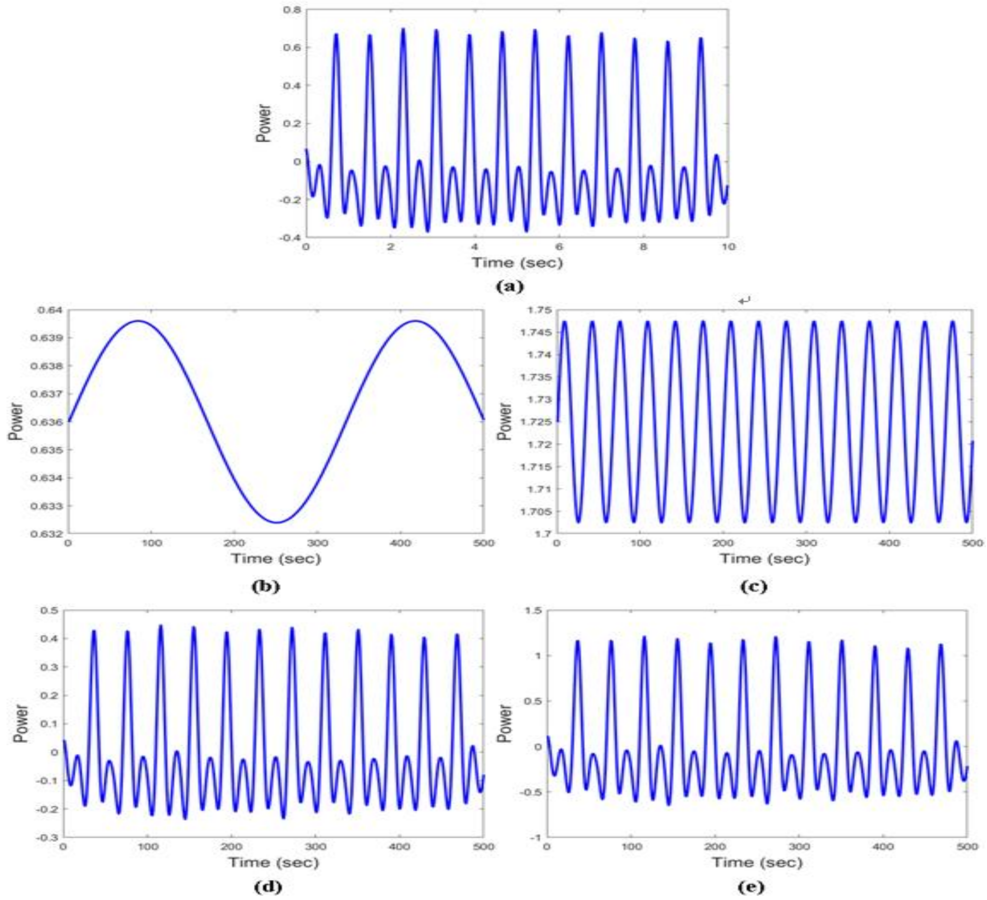
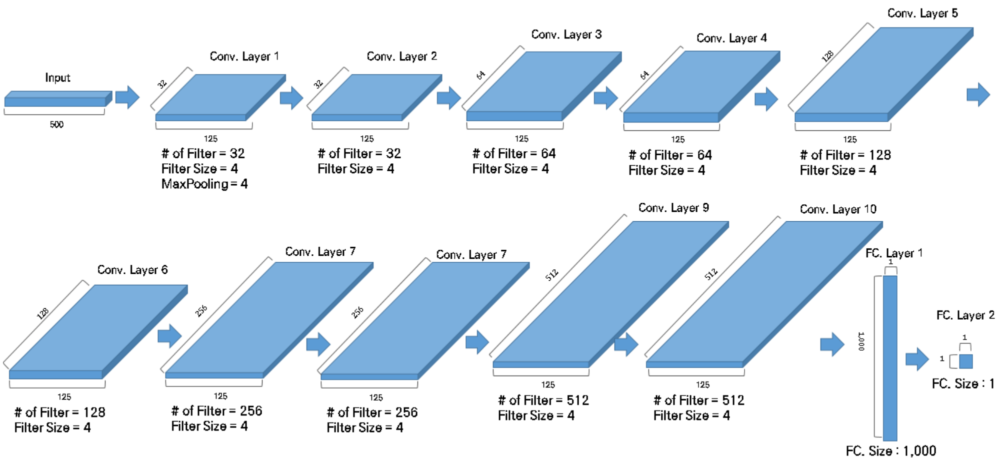


Fig. 3:(a) original PPG signal (b)-(c) Two examples of W(n) (d)-(e) augmented PPG by multiplying W(n) with original PPG signal

In Fig. 3, the Fig. 3(a) is an original abnormal diastolic PPG signal, the Fig. 3(b) is W(n) obtained when P = 0.06, F_c=0.3 Hz, and the Fig. 3(c) is W(n) when P = 0.15 and F_c=3 Hz. By window function W(n) in (1), the amplitude ranges from a minimum of 0.54 to 0.66 to a maximum of 1.35 to 1.65. After obtaining the window function, the W(n) and the original PPG signal are sample-by-sample multiplied. The signal lengths of W(n) and the PPG are the same. By doing this multiplication, a different shape of PPG waveform is made from the original PPG waveform. The window function varies the amplitude of the original PPG signal. By doing this process many times, augmented signals are generated. The resulting signalsare used for training. Fig. 3(d) and 3(e) shows the augmented PPG signals of Fig. 3(b) and 3(c), respectively. Data augmentation is performed repeatedly for the abnormal PPG signals. The number of repetitions is 150, and as a result, the abnormal PPG signal is increased to 150 x 188 = 28,200 signals from the 188 original signals.

5. 1D-CNN

This paper proposes an 1D-CNN (a type of artificial neural network) model for the HR estimation. CNN focuses on a local data, not the whole data. CNN is suitable for the highly correlated input signal. For example, CNN shows excellent performance in image identification. In other words, unlike the other deep natural networks, CNN maintains local information of input signal so that CNN is suitable for the locally correlated input signal such as pictures or correlated 1D signals. As shown in Fig. 3, PPG signal is locally correlated and hence, 1D-CNN is an appropriate choice. Since the PPG signal used in this paper can be viewed as one-dimensional image, 1D-CNN is used for the deep neural network.



5.1. 1D-CNN design

Layer	Number of Parameters
Conv. Layer 1	160
Conv. Layer 2	4,128
Conv. Layer 3	8256
Conv. Layer 4	16,448
Conv. Layer 5	32,896
Conv. Layer 6	65,664

Fig. 4: Proposed 1D-CNN

Table. 1: The number of Parameter of Proposed 1D-CNN

Conv. Layer 7	131,328
Conv. Layer 8	262,400
Conv. Layer 9	524,800
Conv. Layer 10	1,049,088
FC Layer 1	64,001,000
FC Layer 2	1,001
Total	**66,105,105**

The proposed structure of the 1D-CNN is shown in Fig. 4. The 1D-CNN is composed of ten convolutional layers and twodense layers. In ten convolutional layers, the length of the convolution filter is all 4 (1x4). The number of filters is 32 in the 1st and 2nd layers, 64 in the third and fourth layers, 128 in the fifth and sixth layers, 256 in the seventh and eighth layers, and 512 in the ninth and tenth layers. The deeper the layer, the larger the number of convolution filters, which is a natural approach. The Max pooling is used only in the first convolution layer output to reduce the output size to 1/4. The output of the first affine layer is 1,000, and the second affine layer output is 1, which is the estimated HR. Table 1 summarizes the number of learnable parameters of the proposed 1D-CNN at each layer. The first affine layer takes most of the number of parameters, and the total number of parameters is 66,105,105.

6. Training Method

The neural net training is performed with the proposed 1D-CNN and the augmented PPG signals. The number of data used for training is 28,378 normal PPGs and 28,200 augmented abnormal PPGs. The programming tool for training and test are Python 3.7.5 and Keras 2.0.For accelerating the computation, NVIDIA GeForce RTX 3090 is used. The mini-batch size, one of the hyper parameters, is 1,024 and the maximum epoch is 3,000. As an optimization algorithm, AdaGrad is used with initial learning rate of 0.0001.

6.1. Training loss function

The loss function for training is a Mean Square Error (MSE). MSE is the most commonly used loss function in training for regression problem to find the neural network parameters. The MSE is defined by an error or difference between the actual HR value and the predicted HR value. Since MSE squares the HR difference, MSE is more sensitively to outliers than Mean Absolute Errors (MAEs) which is absolute difference between the ground-truth HR and the predicted HR. The MSE is represented as

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{2}$$

6.2. Cross validation

Cross validation is a method of performance verification for fair comparison. When cross validation is used, reliability and fairnesscan be guaranteed. If cross validation is not performed, performance bias can occur because training and validation are performed only on a specific combination of training and validation data sets. By incorporating all possible combination of training and validation data set, fair performance can be obtained. Total PPG signals are divided into 3In this paper, training and validation data split3 to 1 for training and validation, respectively. Thus, there are four combinations of data split. Cross validation is performed by averaging the verification results for all the four combinations. Fig. 5 shows the structure of cross validation.

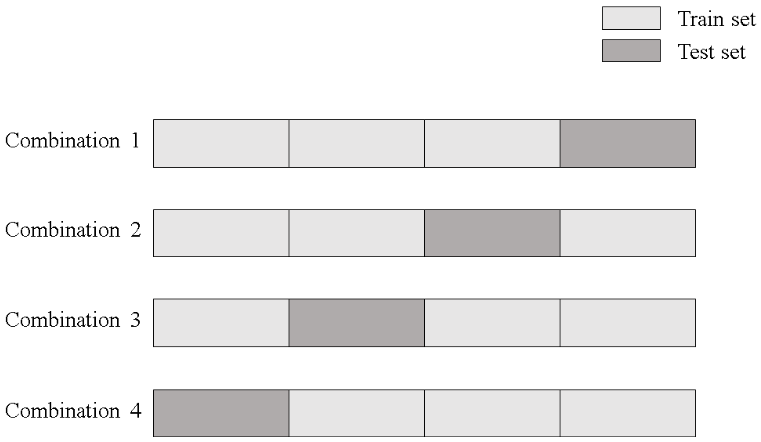


Fig.5: Data Combination for Cross Validation

7. Verification

The accuracy of the HR prediction is measured by MAE. (3) is the formula of MAE.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i) \quad (3)$$

Cross validation performs training and verification for all possible four combinations in parallel. Table 2 is the final verification performance for each combination and averaging all the combinations. The averaged MAE is 1.23 BPM.

Table. 2: Verification results by cross validation

	Data Combination1	Data Combination 2	Data Combination 3	Data Combination 4	**Averaged**
MAE (BPM)	1.19	1.33	1.22	1.21	**1.23**

8. Conclusion

Thispaper proposed an HR estimation technique based on 1D-CNN. The 1D-CNN is composed of ten convolutional layers and two affine layers. Data augmentation is used to make up for the insufficient abnormal PPG signals. The averaged result of performance for four combinations through cross validation showed the MAE of 1.23 BPM. This result indicates that the proposed HR estimator can be an alternative for HR monitor. If the proposed technique is employed in wearable PPG devices, a simple and accurate day-and-night HR monitoring will be possible. However performance verification with other PPG sensors is needed to ensure the generality of the proposed technique. In the future,we will conduct research on this.

9. References

1. Acharya, U.R., Joseph, K.P., Kannathal, N., Lim, C.M., Suri, J.S. (2006). Heart rate variability : a review. Medical and biological engineering and computing, 44(12), 1031-1051.
2. Browne, M.W. (2000). Cross-validation methods. Journal of mathematical psychology, 44(1), 108-132.
3. Christofaro, D.G.D., Casonatto, J., Vanderlei, L.C.M., Cucato, G.G., Dias, R.M.R. (2017). Relationship between resting heart rate, blood pressure and pulse pressure in adolescents. Arquivosbrasileiros de cardiologia, 108(5), 405-410.
4. Eren, L., Ince, T., Kiranyaz, S. (2018). A generic intelligent bearing fault diagnosis system using compact adaptive 1D CNN classifier. Journal of Signal Processing Systems, 91(2), 179-189.
5. Grande, D., Iacoviello, M., Aspromonte, N. (2018). The effects of heart rate control in chronic heart failure with reduced ejection fraction. Heart failure reviews, 23(4), 527-535.
6. Hannun, A.Y., Rajpurkar, P., Haghpanahi, M., Tison, G.H., Bourn, C., Turakhia, M.P., Ng, A.Y. (2019) Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. Nature medicine, 25(1), 65-69.
7. H.G, Kim., E.J,Cheon., D.S, Bai., Y,H, Lee., B.H, Koo. (2018). Stress and heart rate variability: a meta-analysis and review of the literature. Psychiatry investigation, 15(3), 235-245.
8. Lu, G., Yang, F., Taylor, J.A., Stein, J.F. (2009). A comparison of photoplethysmography and ECG recording to analyse heart rate variability in healthy subjects. Journal of medical engineering & technology, 33(8), 634-641.
9. Reimers, A.K., Knapp, G., Reimers, C.D. (2018). Effects of exercise on the resting heart rate: a systematic review and meta-analysis of interventional studies. Journal of clinical medicine, 1(88), 9-20.
10. S, Kwon, J,Hong.,E.K, Choi., B, Lee., C, Baik., E, Lee., E.R, Jeong., B.K, Koo., S, Oh., Y, Yi. (2020). Detection of atrial fibrillation using a ring-type wearable device (CardioTracker) and deep learning analysis of photoplethysmography signals: prospective observational proof-of-concept study. Journal of medical Internet research [Internet], 22(5), e16443.
11. Temko, A. (2017). Accurate heart rate monitoring during physical exercises using PPG. IEEE Transactions on Biomedical Engineering, 64(9), 2016-2024.
12. 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.