1-Dimensional Convolutional Neural Network based Heart Rate Estimation Using Photoplethysmogram Signals

Seong-Hyun. Kim
Hanbat National University, Daejeon, Republic of Korea. E-mail: emzhak2@naver.com

Eui-Rim Jeong*
Hanbat National University, Daejeon, Republic of Korea. E-mail: erjeong@hanbat.ac.kr

Received September 26, 2021; Accepted December 19, 2021
ISSN: 1735-188X
DOI: 10.14704/WEB/V19I1/WEB19303

Abstract

Recently, as the importance of healthcare has increased, researches are being conducted to measure health status in real time. Heart Rate (HR) measurement is one of the important health conditions that measure heart beat rates. HR measurement can be performed using Photoplethysmogram (PPG) or Electrocardiogram (ECG) signals. Since the PPG or ECG signals are different from people to people, conventional HR estimator occasionally results in large errors. To develop a reliable HR estimator, an HR estimation technique using PPG is proposed in this paper, based on a deep learning technique. The proposed HR estimation technique has the following key features. We develop a new artificial neural network which is 1-Dimensional Convolutional Neural Network (1D-CNN) composed of ten convolutional layers and two fully connected layers. To assess the estimation performance, cross validation is used. The training and verification of the proposed 1D-CNN technique are performed on Python 3.7.5 with Keras 2.0 library. The proposed HR estimation technique performs training and verification using field PPG data. Overfitting is prevented by increasing the limited training data by data augmentation. In training, the loss function is the Mean Square Error (MSE), which is commonly used in regression problems. In the verification, the error between the predicted HR and the actual HR is compared using Mean Absolute Error (MAE). As a result of the final performance verification through cross validation, the proposed technique shows an MAE of 1.23 Beats Per Minute (BPM). This results indicate that the proposed technique enables quick and accurate HR estimation with only PPG signals. Therefore, if this technique is applied to medical and wearable devices, the proposed technique can replace the existing HR monitors.

Keywords

Heart Rate, Photoplethysmogram, Convolutional Neural Network, Cross Validation, Mean Square Error.
Introduction

Heart Rate (HR) is the number of heart beats per minute and the measuring unit is Beat Per Minute (BPM) (Acharya et al., 2006). A normal person's HR is between 70 and 80 BPM on average (Christofaro et al., 2017). HR can vary depending on life activities such as physical exercise, sleep, and breathing (Reimers et al., 2018). HR can also vary depending on disease, not on life activity (Kim et al., 2018). For example, in the case of people with arrhythmia, HR suddenly increases (Hannun et al., 2019). In addition, people with high HR have a higher incidence of cardiovascular disease and consequently, the resulting mortality rate is higher than those with low HR (Grande et al., 206). Therefore, it is very important to monitor the HR and check the HR change to determine whether a health problem occurs or not.

HR can be measured using Photoplethysmogram (PPG) or Electrocardiogram (ECG) signals (Lu et al., 2009). Between them, PPG signal is easier to obtain. The PPG measures the intensity of light transmitted from the skin and measures the reflected pulse. PPG signals have the information on the change in blood flow (Temko, 2017). PPG signal has two peaks: systolic peak and diastolic peak (Yousef et al., 2012). Systolic peak is the peak corresponding to the actual pulse and is the output of a direct pressure wave moving from the left ventricle to the surroundings of the body. Another peak, diastolic peak, is the result of reflection of pressure waves by arteries in the lower body. In general, the diastolic peak is very small compared to the systolic peak. However, depending on the person, there are people with a very large diastolic peak. When the diastolic peak is very large, the diastolic peak may be used to estimate HR due to confusion with the systolic peak during training. Since the Diastolic peak is not a peak of the actual pulse, care should be taken not to measure the pulse when measuring HR.

In this paper, we propose a 1-Dimension Convolutional Neural Network (1D-CNN) (Eren et al., 2018; Kwon et al., 2020; Jeong et al., 2020) that estimates HR from PPG signals for reliable HR estimation. The deep learning technique such as 1D-CNN learns from data, and if sufficient and various data are provided, the technique can predict the HR stably and reliably. The PPG used for training consists of both normal signals with small diastolic peaks and abnormal signals with large diastolic peaks. Since the number of data can be insufficient for the abnormal cases, the number of signals is increased through data augmentation. The structure of the proposed 1D-CNN consists of 10 convolutional layers and two Fully Connected (FC) layers. Cross validation is performed to verify the performance of all data, not just some data combination (Browne, 2000). As a result of the final performance evaluation through cross validation, the proposed techniques' MAE is
about 1.23 BPM. If the proposed technique is used for wearable devices, HR can be estimated with high accuracy.

**System Configuration**

![Block Diagram of Proposed HR Estimation Technique](image)

Figure 1. Block Diagram of Proposed HR Estimation Technique

Figure 1 shows the system configuration diagram proposed in this paper. First, PPG data to be used for training for cross validation are divided into four combinations. Among the divided data, the number of PPG data of abnormally large diastolic peaks increases through data augmentation. By using all the augmented abnormal PPGs and normal PPGs, the proposed 1D-CNN is trained. The cross validation enable to avoid the performance variation due to the combinations of the training and test data sets. The average performance of all the combinations can show a reliable result.

**PPG Data Set**

In this paper, HR estimation is performed using PPG signals. PPG is one of the methods of measuring heartbeat. PPG measures pulse by measuring the degree of transmission of light between the light source and the light receiver. Among the components in the path through which light is transmitted, the components that can absorb light are the amount of skin, tissue, and blood. In this case, it is an ingredient that does not change except for the amount of blood. Therefore, it is possible to know the amount of blood that changes by heartbeat, and through this, heartbeat measurement is possible. The data used for training and verification are total 28,378 normal PPGs and 188 abnormal PPGs. The PPG signal is sampled with 50 Hz clock frequency for 10 seconds, and consequently, the length of one PPG data is $50 \times 10 = 500$ samples.
Problems by Large Diastolic Peaks

The peak corresponding to the actual pulse of PPG is a systolic peak. Systolic peak is the output of a direct pressure wave moving from the left ventricle to the surroundings of the body. In addition to the systolic peak, there is a diastolic peak in PPG. Diastolic peak is the result of reflection of pressure waves by arteries in the lower body. Figure 2 shows the cases for a normal PPG and an abnormal PPG where diastolic peaks are large, respectively.

![Figure 2: Two case of PPG: normal PPG (left) and abnormal PPG (right)](image)

When the diastolic peak is large as shown in the PPG on the right side of Figure 2, not only the actual peak systolic peak but also the diastolic peak can be observed on the PPG signal. In this case, a risk exists that the HR is estimated twice the actual HR. Therefore, training data for cases where the diastolic peak is large is very important. To this end, data augmentation is done to increasing the number of data when the diastolic peak is large. Data augmentation solves the lack of data when the diastolic peak is large.

Data Augmentation

Data augmentation is performed before training the 1D-CNN. The objective of data augmentation is to increase the number of the abnormal PPG signals to make the numbers of normal PPGs and abnormal PPGs similar. Data augmentation in image processing field is a technique that increases the number of images by making minor changes to the input image. Data augmentation is a particularly effective tool when there is only a few data. For data augmentation, a window function is defined. The window used for data augmentation is expressed in (1).

\[
W(n) = P \times (\sin(2 \times \pi \times F_c \times n + Offset) + 10) \text{ for } n = 0, ..., 499
\]  

(1)
In this case, the P is randomly chosen in 0.06 to 0.15, the $F_c$ is in 0.3 to 3, and the Offset is in 0.01 to 1. Some example shapes of $W(n)$ that is made through (1) are shown in Figure 3.

![Figure 3](image.png)

**Figure 3.** (a) original PPG signal (b)-(c) Two examples of $W(n)$ (d)-(e) augmented PPG by $W(n)$
In Figure 3, the Figure 3(a) is an original abnormal PPG signal, the Figure 3(b) is W(n) obtained when $P = 0.06$, $F_c = 0.3$ Hz, and the Figure 3(c) is W(n) when $P = 0.15$ and $F_c = 3$ Hz. When (1) is performed, the amplitude of the window ranges from a minimum of 0.54 to 0.66 and a maximum of 1.35 to 1.65. After obtaining the window function, the original PPG signal is sample-by-sample multiplied with W(n). The signal lengths of PPG and window are the same. By doing this process, a different shape of PPG waveform is generated from the original PPG waveform, and the resulting PPG signal is used for training. Figure 3(d) and 3(e) shows the augmented PPG signals by Figure 3(b) and 3(c), respectively. Data augmentation is performed by repeatedly taking the augmentation for the abnormal PPG. The number of repetitions is set to 150, and as a result, the abnormal PPG signal is expanded to $150 \times 188 = 28,200$ signals from the 188 original signals. The figure 3(d) and (e) are examples of data augmentation.

**1D-CNN Design for HR Estimation**

A technique for estimating HR using 1D-CNN (a type of artificial neural network) is introduced. In addition, CNN looks at a part of the input data, not the whole, to determine the relationship between one pixel of the image and surrounding pixels. Therefore, CNN shows effective performance in image analysis. In other words, unlike the other deep natural networks, CNN maintains spatial information of input data so that CNN is suitable for the locally correlated input data such as pictures or temperatures over time. As shown in Figure 3, PPG signal is locally correlated and 1D-CNN is a natural choice. Since the configuration of the data used in this paper can be considered as one-dimensional image, 1D-CNN is used for the artificial neural network.

![Figure 4. Proposed 1D-CNN](http://www.webology.org)
### Table 1: The number of Parameter of Proposed 1D-CNN

<table>
<thead>
<tr>
<th>Layer</th>
<th>Number of Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conv. Layer 1</td>
<td>160</td>
</tr>
<tr>
<td>Conv. Layer 2</td>
<td>4,128</td>
</tr>
<tr>
<td>Conv. Layer 3</td>
<td>8,256</td>
</tr>
<tr>
<td>Conv. Layer 4</td>
<td>16,448</td>
</tr>
<tr>
<td>Conv. Layer 5</td>
<td>32,896</td>
</tr>
<tr>
<td>Conv. Layer 6</td>
<td>65,664</td>
</tr>
<tr>
<td>Conv. Layer 7</td>
<td>131,328</td>
</tr>
<tr>
<td>Conv. Layer 8</td>
<td>262,400</td>
</tr>
<tr>
<td>Conv. Layer 9</td>
<td>524,800</td>
</tr>
<tr>
<td>Conv. Layer 10</td>
<td>1,049,088</td>
</tr>
<tr>
<td>FC Layer 1</td>
<td>64,001,000</td>
</tr>
<tr>
<td>FC Layer 2</td>
<td>1,001</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>66,105,105</strong></td>
</tr>
</tbody>
</table>

The structure of the designed 1D-CNN is shown in Figure 4. The proposed 1D-CNN consists of 10 convolutional layers and 2 fully connected layers. In 10 convolutional layers, the length of the convolution filter is all 4. The number of filters is 32 in the 1st and 2nd layers, 64 in the 3rd and 4th layers, 128 in the 5th and 6th layers, 256 in the 7th and 8th layers, and 512 in the 9th and 10th layers. The deeper the layer, the larger the number of convolution filters. The Max pooling is used in the first convolution layer only to reduce the output size to 1/4. The output of the first fully connected layer is 1x1,000, and the second fully connected layer output is 1x1, which is the estimated HR value, the final output. Table 1 summarizes the learnable parameters of the proposed 1D-CNN at each layer. The first fully connected layer occupies most of the number of parameters, and the total number of parameters is 66,105,105.

### 1D-CNN Training

The training is carried out with the designed 1D-CNN and the augmented PPG signals. The data used for training are 28,378 normal PPGs and 28,200 augmented abnormal PPGs. The tool used for training and test are Python 3.7.5 with Keras 2.0 library, and computations are performed using the NVIDIA GeForce RTX 3080. The mini-batch size, which is a hyper parameter used for training, is 1,024 and the maximum epoch is 3,000. As an optimization algorithm, AdaGrad is used, and the initial learning rate is 0.0001.

The loss function used for training is a Mean Square Error (MSE). MSE is the most commonly used loss function in training in regression problem during deep learning. The MSE determines an error based on the difference between the actual HR value and the predicted HR value. Since MSE squares the error value, it acts more sensitively to outliers.
than Mean Absolute Errors (MAEs) that do not squared the error value. The formula of MSE is shown in (2).

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

(2)

**Cross Validation**

Cross validation is a method of making use of all data for training and verification. When cross validation is used, reliability and validity may be given to the verification result. If cross validation is not performed, data bias occurs in training and verification because training and verification are performed only on a specific combination of training and validation data sets. In this paper, training and verification data divided by a ratio of 3:1, and hence, there are four combinations of data split. Cross validation is performed by obtaining the average of the verification results of the four combinations of datasets. Figure 5 is the structure of cross validation.

![Data Combination for Cross Validation](image)

**Performance of Proposed Technique**

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

\[
MAE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)
\]  

(3)

Cross validation performs training and verification for four combinations in parallel. By obtaining the average of the verification results performed in parallel, verification results for all data are obtained. Table 2 is the final verification performance obtained by calculating the verification results and averages for each combination.
Cross validation performs training and verification for four combinations in parallel. By obtaining the average of the verification results performed in parallel, verification results for all data are obtained. Table 2 is the final verification performance obtained by calculating the verification results and averages for each combination.

### Conclusion

In this paper, we propose an HR estimation technique using 1D-CNN. The proposed 1D-CNN consists of 10 convolutional layers and two fully connected layers. Data augmentation is used to make up for insufficient data. It gives validity and reliability to all data by using cross validation during training and verification. As an average result of performance for four combinations through cross validation, the MAE of the proposed technique shows 1.23 BPM. This result confirms that the proposed 1D-CNN HR estimator can be used as an accurate HR monitor. If the proposed technique is applied to wearable PPG devices, it will be of great help in 24-hour, fast and accurate HR monitoring. However, in order to prove the generality of the proposed technique, further studies are needed to determine whether it shows reliable HR values for other PPG sensor signals. We will confirm this generality through additional research in the future.

### References


