



Spectro-Temporal Incorporated Deep Neural Network with Residual Connections for Heart Rate-Based Classification of Photoplethysmogram Signal

P. Deepak Franklin^{1*}, Dr.M. Krishnamoorthi²

Abstract

This paper focused on estimating Blood Pressure (BP) and Heart Beat Rate (HBR) to identify heart diseases using a photoplethysmogram (PPG). Blood pressure is one of the major sources to identify hypertension and criticality conditions. Thus, regular monitoring of the BP regularly is most important. But most patients have an aversion toward cuff-based medical devices used only at rest time. A Spectro-Temporal Deep Neural Network (STDNN) with a residual connection model is proposed to analyze and classify the heart rate. STDNN analyses the PPG data to estimate BP and heart rate as a potential solution investigated. The proposed STDNN model learns the entire PPG signal data and measures the spatiotemporal features using the deep convolution neural network. Parallely, it measures the set of all features and their variation related to heart rate. It is too difficult to predict heart diseases only from heart rate variation but also needs the blood volume. The entire process of STDNN has been divided into stages, such as preprocessing, clustering, and classifying the data regarding heart diseases. Initially, the derivative methods are applied to the PPG signal, get the first-order and second-order derivative values for calculating the variations of the HR values, and feed as input to the STDNN model. The experimental results show that the proposed STDNN model outperforms the other existing models in terms of prediction accuracy. The testing process is carried out on various datasets to obtain the best performance.

Key Words: PPG Singal Processing, Deep Learning, Deep Neural Network, Spatio-Temporal Model, Feature Extraction, Classification, Prediction, Heart Diseases.

DOI Number: 10.14704/nq.2022.20.8.NQ44051 NeuroQuantology 2022; 20(8):446-458

449

Introduction

Heart rate calculation is the number of times heartbeats per minute. If the heartbeat ranges from 60 to 80 per minute, it is normal in an average person. Heart rate monitoring gives personal details about their fitness level and their medical conditions. At the same time they do any physical exercises, the heart rate booms, which results in burning fat. There are several methods available to determine the heart rate of a person. They are electrocardiogram (ECG) and Photoplethysmographic (PPG). Calculating the

heart rate by ECG is accurate, and the heart's readings are directly recorded using the electrodes present in ECG. The difference between ECG and PPG is thus ECG requires the patient to wear the electrode, but it is not necessary for PPG. In PPG, only two LEDs are needed, and this LED is placed on the patient's wrist.

In PPG, the signals are obtained using a pulse oximeter preinstalled in the smart devices attached to the patient body.

Corresponding author: P. Deepak Franklin

Address: ^{1*}Assistant Professor, Department of EEE, Sriram Engineering College, Perumalpattu; ²Professor, Department of CSE, Dr. NGP Institute of Technology, Coimbatore.



E-mail:¹deepftr2020@gmail.com, frankece@gmail.com;²krishna_bit@yahoo.co.in

In PPG, the heart rate calculates the light intensity changes after placing the diodes in the patient's wrist. The two light diodes on the person's wrist lighten on different frequencies. The light signals get deflected after hitting the oxygenated blood, thus calculating the amount of reflection in the heartbeat rate. Under normal circumstances, the frequency from the diodes gets obtained by calculating a simple time-domain signal analysis. The main challenges in retrieving the heart rate from PPG are noisy and distorted signals. It is called Motion Artifacts (MA). Motion artifacts are used in retrieving the heart signals in both frequency domain analyses.

Many signal processing techniques have been designed to identify the motion artifacts and remove the noise level from distorted signals. Adaptive Noise Cancellation techniques are used in noise cancellation procedures [Zhilin Zhang et al. (2015), Zhilin Zhang (2015)]. Some previous methods calculated the heart rate by measuring the pulse transmit time (PPT) [X. Ding et al. (2017)]. At present deep learning techniques are used in the computational process of solving the classification problem. Deep learning techniques are used in the automatic learning process and adapt to the optimal features and processing of real signals. The general framework of deep learning includes deep belief networks (DBN), stack encoders, convolutional neural networks, etc. Deep learning technologies are used in image recognition, speech training, and other object identification process. Deep learning is used to train a massive dataset in the biomedical engineering field. The training data includes heart rate data and brain waves and is even used in tumor recognition. In such cases, signals from PPG are trained under deep learning networks.

PPG signal contains many physiological signals since heart rate has more information about the human body. Presently, PPG replaces all other cardiovascular processing [G. Martínez et al. (2016)]. While collecting the PPG images, many signal processing methods are applied to obtain a clear signal with accurate results. In such cases, adaptive filtering methods separate the heart component from the motion artifact and noisy signal levels.

Many types of research works are going on in the representation of signals in the frequency-time domain. The wavelet transformation method has been used to estimate the heart rate and denoise

the PPG signals. Some earlier research works used the time-frequency spectrum to analyze the PPG signal, separating the frequency levels from the PPG signals. Each method has its limitation. The frequency spectrum analysis method relies on frequency emission, where the noisy signals cannot be identified. While using wavelet transmission, the results are inaccurate because the wavelet transformation uses standard procedures and does not tend to calculate the frequency signals from PPG. The wavelet transform is applied to the PPG signals and trained using convolutional neural networks.

An end-to-end process is suggested while dealing with PPG analysis. Most deep learning techniques use the time-frequency spectrum method to process the PPG signals and produce accurate heart results. Convolutional neural networks can calculate continuous periodic effects [Hammerla et al. (2016)]. In PPG-signal ambulatory atrial fibrillation, a 50-layer CNN is used to measure the presence or absence of atrial fibrillation [Voisin et al. (2018)]. The main disadvantage of PPG signals is frequency separation is used to obtain an accurate heartbeat rate. PPG is used in various medical applications, such as measuring the oxygenated blood and blood pressure and even assessing cardiac output. It also detects peripheral vascular disease. PPG is an optimal and straightforward tool for detecting changes in blood volume. PPG is considered non-invasive as it measures only through the skin surface.

This paper proposes an efficient, deep learning method for detecting the heart rate from the PPG signals based on the Spatio-temporal features. The paper's overall process involves four stages: signal preprocessing, spectral estimation, temporal estimation, and signal reconstruction. The obtained spectral and temporal information is fed to the CNN model for training.

Related Works

Biswas et al. (2019) discussed the state-of-the-art methods used for PPG-based heart signal analysis to estimate the heart rate. Biswas et al. [2019] generated a review based on the PPG heart signal deduction. More extensive filtering methods are necessary for obtaining the heartbeat rates. Detecting heart rate is used in finding the balance between stress and recovery. The existing model's limitations are the periodic lag in heart rate



deduction—the filtering methods [B. Lee et al. (2010), M.R. Ram (2012)] is applicable to separate the motion artifact. This model is carried out by traditional signal processing techniques with finite and infinite pulse responses. It uses high pass filtering moving average filtering, and wavelet transformation to reduce the motion artifact. The limitation in this model is unequal frequency. The unequal frequency provides inaccurate results. The other noise signals from the heart are obtained from heart signals from PPG, and they are carried out at different time series [B.X. Kim et al. (2006), X. Sun (2012)]. Recent research mainly focuses on signal representations which are highly useful in heart beat monitoring. A typical PPG spectrum detects the periodic heart rate and periodic physical activities. Deep learning techniques are already applied in different time series, such as human activity recognition or gait parameter extraction [M. Zheng et al (2018), J. Hannink et al. (2017)]. Deep learning is expensive, and training the data models requires more data. Deep learning requires more expensive GPUs. Deep learning and the optimal machine learning techniques are compared in extracting the data from human activity. They show that if sufficient data is available, both the networks can extract the human heart rate and other physical activities more accurately.

Jindal et al. [2016] performed an analysis technique with deep neural networks and the results obtained are mostly accurate. However, some cases find inaccurate results when subjected to vigorous physical activities. This method also requires many unique features as input for the training process, and this approach is not an end-to-end learning technique. The author used a preprocessing time-

series signal as input for convolutional neural networks and the LSTM model in performing the PPG-based heart rate estimation [Biswas et al. (2019)].

Recent work related to PPG is replacing arterial blood pressure with PPG signals. The feature extraction processes are used in extracting the features from the PPG signals. However, there occurs a limitation in feature extraction, such as the heart rate and their exact features must be extracted correctly [M. Elgendi et al. (2013), M. Elgendi (2014)]. The equipment that ECG and PPG use must be of high quality in case of feature extraction, and the most required feature is that optimal filtering is required [M. Elgendi (2016), M. Elgendi et al. (2015)]. ECG data are more sensitive in nature, and sometimes high voltage data is needed in data extraction. Though ECG is accurate in results, PPG is widely used to get more accuracy in data classification. The main disadvantage of PPG is the placement problem. The PPG electrodes must be placed in the exact location so that the results are accurate. If the placement varies, the result will be less accurate. In this proposed model, the filtering models are less accurate because of the poor data obtained. It is thus resolved by using a high-quality filtering package.

Spectro Temporal Based Deep CNN

The models mentioned above have their positives and limitation. The major limitation of all models is that the models require more data for filtering to remove the motion artifacts. More data is obtained in this proposed model, and the quality data package is used for filtering for removing the motion artifacts in heart rate deduction.



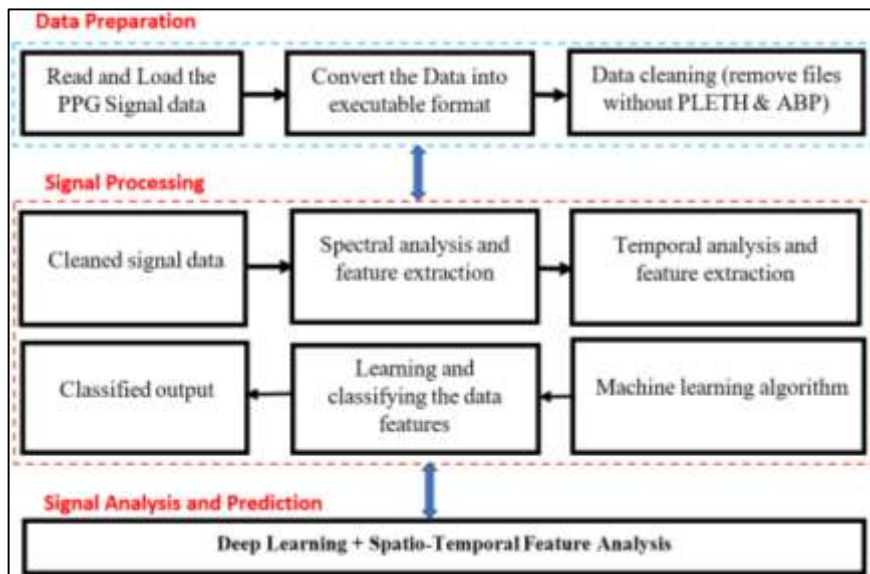


Figure 1. The architecture of the Proposed Model

The current goal is to identify the spectro-temporal characteristics of PPG signals such that deep convolution neural networks (CNNs) can classify it.

The proposed comprehensive architecture of input (PPG) to outcome (expected label) is shown in Figure-2.

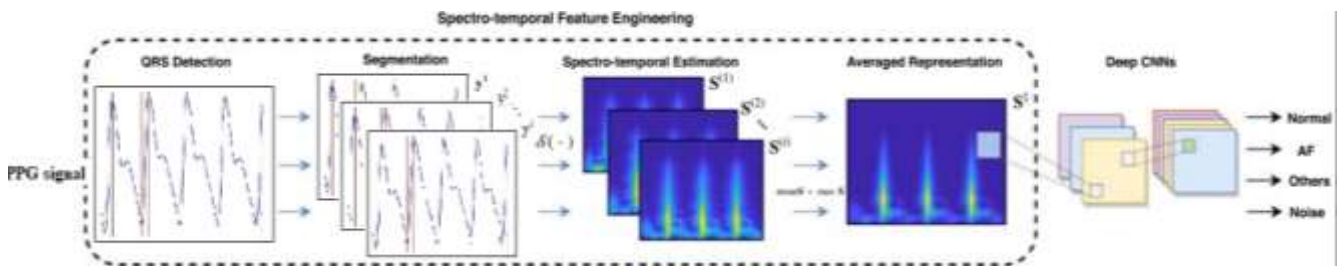


Figure 2. Spectro-Temporal Feature Extraction

The original ECG signal is separated into fixed-length sections and oriented by its center R maxima in the initial phase, QRS identification, and ECG segmentation. Subsequently, the expression $[S]_{j,k} = \sqrt{\hat{a}_j^2(t_k) + \hat{b}_j^2(t_k)}$ is used to determine the spectro-temporal matrix for every sector. Once the data matrices have been normalized and aggregated, a fixed-length spectro-temporal feature matrix is produced. The 2D feature set (spectro-temporal image) is put into a deep CNN in the last phase to do classification.

The feature extraction process that has segmented and averaged strategies (shown in the dashed region of Figure-2) is justified by three different arguments. Initially, it can create a fixed-length spectro-temporal feature matrix to address the issue of PPG data with the varying lengths of CNN. Furthermore, it could collect sufficient data from a PPG to allow CNN's to classify it. For illustration, sharp edges are predicted concerning the QRS complexes available in the feature matrix to obtain

usual heart rhythms. Due to that, the focused R maxima in every section are properly oriented, followed by the averaging process. A noisy environment in the spectro-temporal images is predicted for AF rhythms due to the R-R intervals changes. In those sections, there is hardly any lucid zone for the respective QRS complex. In other classes, researchers predict the various patterns observed in the spectro-temporal images depending on the actual arrhythmia. Eventually, the phases such as segmentation and averaging are primarily employed to reduce the noisy environment present in the PPG data. In the subsequent sections, various steps involved in feature extraction are explained in detail.

In this study, we modify the Pan-Tompkins method for QRS identification. Burst noise can readily be misinterpreted by the original Pan Tompkins method [1], which is also susceptible to a noise having an R peak. We gently alter the earliest method to verify the quantity of identified R maxima successfully. If the amount is below a cut-off point, reject the identified R maxima and their



surrounding specimens in the PPG signal. It is done before applying the Pan-Tompkins technique to the remaining signals to overcome this constraint somewhat. In this approach, our techniques can manage the rare situations of high-amplitude burst noise.

The next step is segmentation, in which the fixed-length PPG segments are extracted from the original signal such that each segment potentially covers three QRS complexes. The following process explains the segmentation. If $y = [y_1 y_2 \dots y_N]^T \in \mathbb{R}^N$ denotes the actual PPG signal, whereas the term $\bar{p}_i \in \{1, 2, \dots, N\}$ denotes the location of i^{th} R maxima in y . The term $\bar{P} = [\bar{P}_1 \bar{P}_2 \dots \bar{P}_D]^T$ denotes the location of every R maxima, and the term D denotes the overall quantity of R maxima contained in y . Next, to retrieve the features in the $D-2$ PPG sections, every section in $\bar{p}_i, i \in \{2, \dots, D-1\}$ is combined with the section of y . It is performed to wrap other three QRS complexes successfully. To accomplish it, preceding and following the \bar{p}_i , various β specimens are gathered. After the process mentioned above, the PPG section linked with the i^{th} R maxima could be successfully extracted after y . It is represented as $y^{(i)} = [y_{\bar{p}_i-\beta} \dots y_{\bar{p}_i} \dots y_{\bar{p}_i+\beta}]^T$ whereas the respective PPG section is represented as $S^{(i)} \in \mathbb{R}^{M \times (2\beta+1)}$, t the term M denotes the frequency, and the term $2\beta+1$ denotes the time phases.

The dimension of the matrix S is determined by those two factors, namely M and $2\beta+1$. The selection of β factor is crucial since it controls the outcome's length and the amount of mean taken into account. Generally, periods must comprise at least three QRS complexes for strong proof of R-R.

Algorithm for Averaging Description

Input: The input signal, $y = [y_1 y_2 \dots y_N]^T$

Output: The feature of Spectro-Temporal feature $S^\ddagger \in \mathbb{R}^{M \times (2\beta+1)}$

1. Implement Pan-Tompkins over y and get \bar{p}
2. $D = \text{size}(\bar{p})$
3. **If** $D \leq \delta$ **then**
4. $\bar{y} = y$
5. $\bar{y}_{i-\alpha:i+\alpha} = 0$ for all $i \in \bar{p}$
6. Implement Pan-Tompkins on \bar{y} , and get novel \bar{p}, D
7. **End if**
8. **for all** $1 < i < D$ **in** \bar{p} **do**

9. Implement Spectro-temporal prediction on $y^{(i)}$ to obtain $S^{(i)}$

10. **end for**

11. **return** $S^\ddagger = \frac{\sum_{i=2}^{D-1} S^{(i)}}{D-2} \circ \max_{2 \leq i \leq D-1} S^{(i)}$

By averaging on all spectro-temporal feature matrices and scaling by its maximal masks, the spectro-temporal feature matrix S^\ddagger is produced:

$$S^\ddagger = \frac{\sum_{i=2}^{D-1} S^{(i)}}{D-2} \circ \max_{2 \leq i \leq D-1} S^{(i)}$$

The inclusion of a max function in the expression above is justified since it may, to some degree, aid in maintaining minute aspects of spectro-temporal data that would have been destroyed in segment-by-segment averaging as well as normalizing the data.

Classification

Deep learning methods, particularly convolutional neural networks, have seen considerable progress in identification and classification applications over the last decade. The development of CNNs for 2D image uses highly successful than that of 1D CNNs designs. The objective is to use time-varying spectra and sophisticated CNNs for AF classification.

The data obtained throughout training, particularly the gradient, may disintegrate if the structure is profound and may possess many layers, which would be a "vanishing gradient" problem. It also affects a majority of the existing network designs [2]. This fundamental issue can be resolved in several straightforward methods, such as pre-training, residual connections, or carefully chosen activation functions (for example, one shouldn't connect ReLu before batch normalization).

The network that obtained the best paper position in 2017 that belongs to CVPR is Densely linked Convolutional Networks, usually termed as DenseNet [3]. It can deliver very good efficiency by neglecting degradation and over-fitting issues even if it possesses a stack of several layers. Deep residual networks (ResNets) [4] and DenseNets are improved analogs of each other, with the earlier introducing direct connections on each two and previous layers in a dense frame as opposed to just neighboring layers, as seen in Figure 3. The other merits of this network include the re-utilization of features [3].



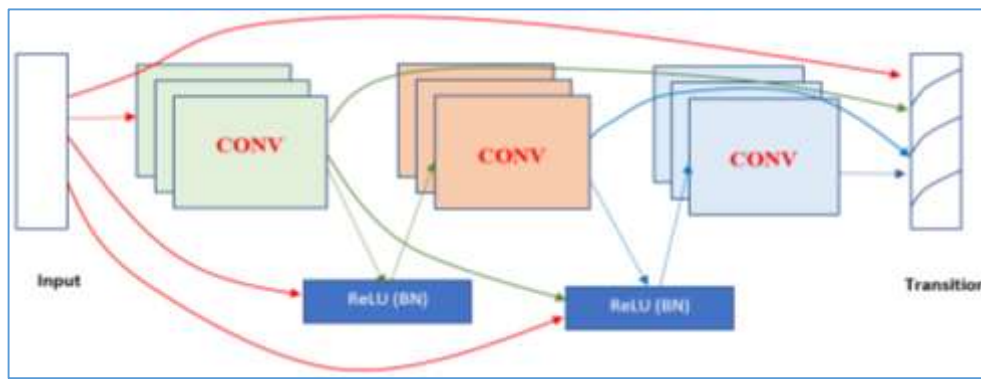


Figure 3. Proposed CNN with Dense Layer

Let us consider the L-layers model and the input image U_0 , then the outcome of the l^{th} -layer is given by:

$$U_l = H_l^{Res}(U_{l-1}) + U_{l-1}$$

$$U_l = H_l^{Den}([U_0 U_1 \dots U_{l-1}])$$

The terms H_l^{Res} and H_l^{Den} denotes the operation of the layers, including convolution, activation, or batch-normalization present in the ResNet and DensNet networks. The term U_l denotes the outcome at the l^{th} layer. The network applied in this work is termed Dense18+ and it is distinct from the actual proposed method. In this, max and average global pooling is concatenated along with the fully connected layer and illustrated in Table-1. The input dimension eliminates the down-sampling max pooling layer in this specific application. Every block in the network possesses 3×3 convolutional layers, and the increasing range equals 48.

Table1. Proposed CNN Structure

Layer Name	Structure	Output Size
Input Convolution Stride 1	Input 7×7 conv (50,50,64)	(50,50,1)
Dense Block 1 Transition 1 2×2 ave pool	$\begin{bmatrix} 1 \times 1 & conv \\ 3 \times 3 & conv \end{bmatrix} \times 4$ 1×1 conv(25,25,128)	(50,50,256)
Dense Block 2 Transition 2 2×2 ave pool	$\begin{bmatrix} 1 \times 1 & conv \\ 3 \times 3 & conv \end{bmatrix} \times 4$ 1×1 conv(12,12,160)	(25, 25, 320)
Dense Block 3 Transition 32×2 ave pool	$\begin{bmatrix} 1 \times 1 & conv \\ 3 \times 3 & conv \end{bmatrix} \times 4$ 1×1 conv(6,6,176)	(12,12,352)
Dense Block 4 Pooling	$\begin{bmatrix} 1 \times 1 & conv \\ 3 \times 3 & conv \end{bmatrix} \times 4$	(6,6,368)
Concatenate	$\begin{bmatrix} global & ave \\ global & max \end{bmatrix}$ concat	(736)
Fully Connected (Softmax)	4 classes	(4)

The proposed spectro-temporal using CNN is assessed and evaluated by calculating the F1-Score, using the following equation. We have carried several studies on the PPG dataset to assess the efficiency of the suggested strategies.

$$F_1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$$

The PPG signals are used as input to classify the patients' condition in terms of heart rate. Based on the classes given in JNC7, BP level is classified as usual, pre-hyper, and hypertension. An illustrative experiment is carried out to classify various categories and compare them to organize them. All the signals were preprocessed, modeled, and evaluated by implementing them in MATLAB-R2018 software. The required CNN toolbox for GoogLeNet is downloaded from the internet and loaded into the software. From the experiment, the PPG signals are analyzed regarding various factors that influence real-time healthcare.

Experimental Results and Discussion

The experiment collects 100 to 500 PPG signals from Appollo Hospital, KK. Nagar, Madurai. The advantage of the proposed method is that it is applied to the PPG signal data to obtain the spectro-temporal and time-frequency information. The obtained results are compared with one another and with the earlier research methods. The CWT technique is used for signal decomposition, whereas the earlier methods have used STFT, CWT, BurgAR, and OscKS. A sample of the reconstructed PPG signal that contains the basic frequency level, second temporal, and third temporal level are shown in Figure-4.

The period of the segmentation window should be a wise choice when the signal is being processed. Sufficient duration should be provided to foresee a hike in the frequency resolution. The cardiac rhythmic variation produces an aperiodic PPG signal over a long window. Hence, multiple frequency spectrums are present in the long window within the considerable variation range in cardiac rhythm. The proposed method's window length is fixed to 8s and is optimal for the heartbeat rate. It, in turn, enables us to possess a prediction rate without an aperiodic signal.



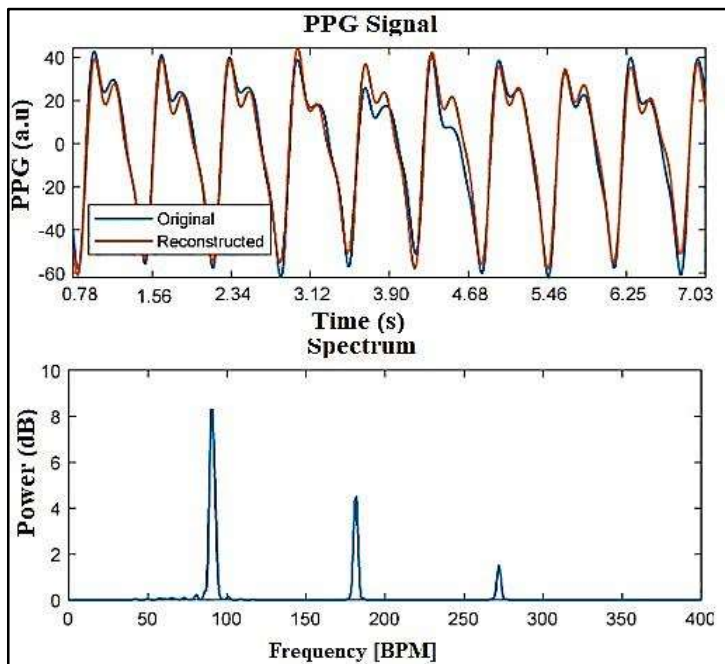


Figure 4. Sample of a reconstructed PPG Signal

The proposed method obtains the recovered PPG signal using a spectro-temporal algorithm, even in noise. More precisely, the proposed algorithm inherits several crucial attributes of the earlier methods and produces a fundamental property for the periodic signals. It is primarily divided into four important phases: signal decomposition, spectral estimation, temporal relationship estimation, and reconstruction. Using the initial two phases, an estimation is attained, which is an 8-s signal component. It is efficiently utilized to extract cardiac functionality. Eventually, the reconstructed signal is obtained. A sample decomposed PPG signal is shown in Figure-8. Sample temporal relation estimation is shown in Figure-6. The green peaks denote the signal's fundamental, second and third temporal, spectral range. The red peaks

denote the signal's identified fundamental, second, and third temporal spectral range. The term **Fund_{cand}** denotes the group of peaks with 10 BPM less frequency isolated from the mean Heart Rate (HR) value. **Second_{cand}** denotes the group of peaks with the frequency nearer to the meanHR (two times). **Third_{cand}** denotes the group of peaks with the frequency nearer to the meanHR (three times). Using the earlier width estimated, a narrow band pass filter is used over the three components recognized. In this phase, minimal-order filters with 60dB stop band attenuation were employed. The filter provides the latency which is required at this phase. In the end, the processed PPG signals were summed up and recovered. The reconstructed PPG signal is shown in Figure-8.



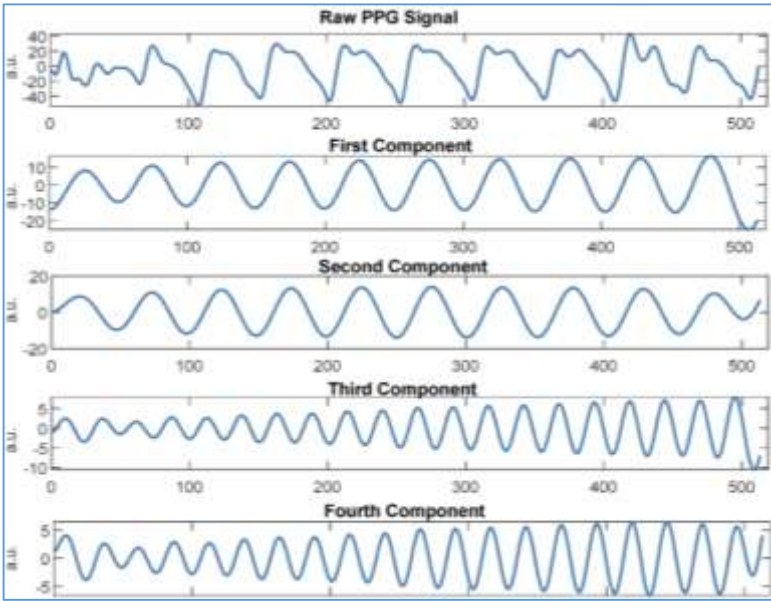


Figure 5. Sample decomposed PPG signal

456

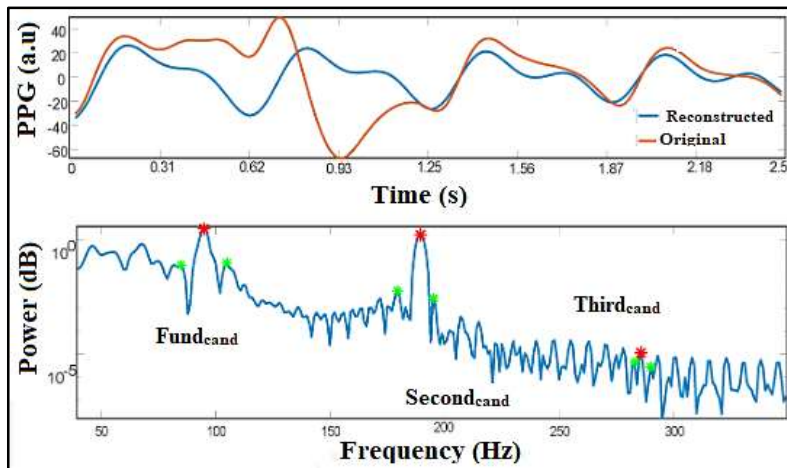


Figure6. Sample temporal relation estimation



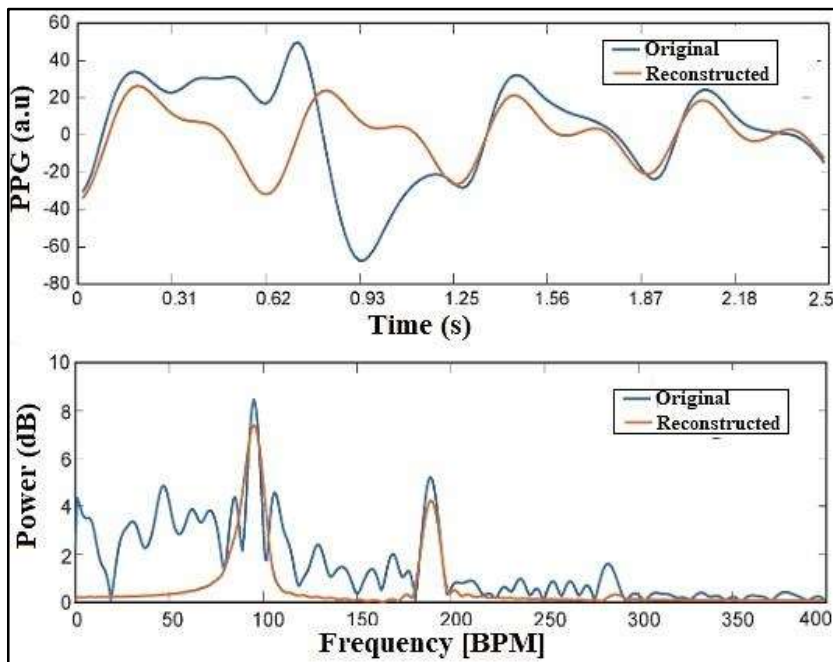


Figure 7. Sample of a noisy reconstructed PPG signal

Spectro-Temporal is mainly employed to reconstruct the PPG signal's actual and full waveform structure in the presence of noise. Its overall purpose is to enable the whole PPG waveform to be made accessible by the methods

that identify its distinctive feature, independent of the wave motions, to enable the subsequent segmentation technique to operate. The primary fiducial locations from which PPG indicators are derived are shown in Figure 8.

457

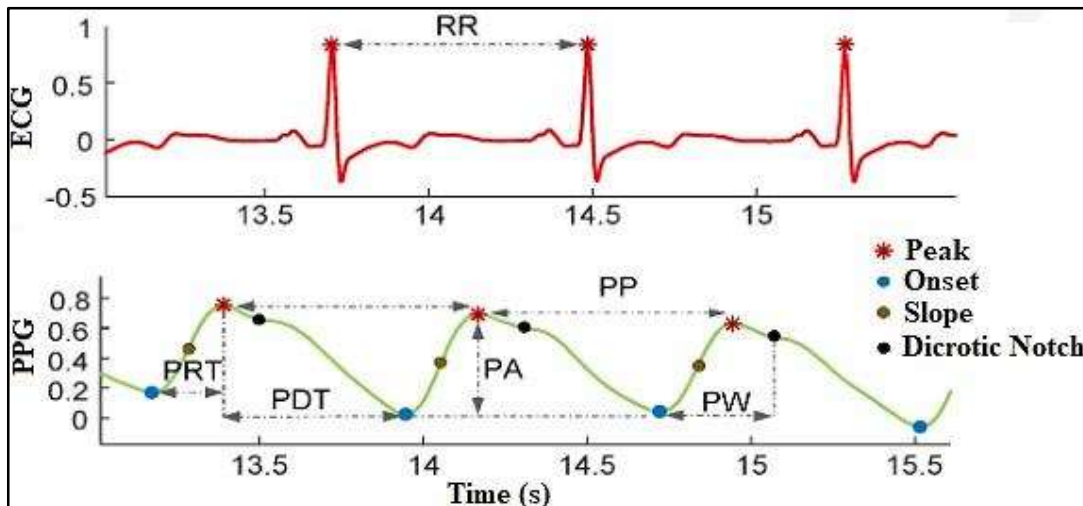


Figure 8. Fiducial locations extracted from ECG and PPG signals

An authenticated standard PPG signal is required to evaluate the efficiency and consistency of Spectro-Temporal, but no such information is currently accessible. As a result, we used a normal PPG signal that we purposefully contaminated with artificial noise to assess the algorithm's effectiveness. As a

result, we have standard fiducial spots that were created by executing the delineation technique on the pure signal to assess the precision of the delineation, as well as a standard PPG signal that can be utilized as ground truth for calculating the reconstruction error.



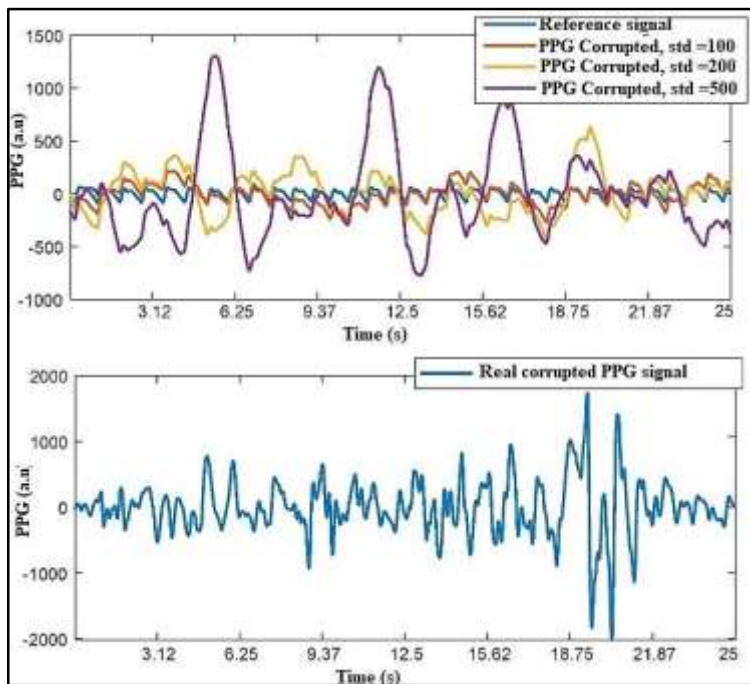


Figure 9. Standard and corrupted PPG Signals

As shown in Figure-8, Although Spectro-Temporal can extract extra cardiovascular data in these circumstances, we can only evaluate the HR identification because there is no other ground truth data source. With four alternative constants for sigma, notably 0 (standard signal), 100, 200, and 500, Figure-9 illustrates a sample of a section

of the reference PPG signal corrupted by our artificial noise.

458

Additionally, it illustrates how noise normally affects a PPG signal. The standard deviation of the shown signal is Figure-9 is 417. With frequencies in the 400–500 bands, experimental data analysis demonstrates that we can get a real noise threshold through light physical exercise.

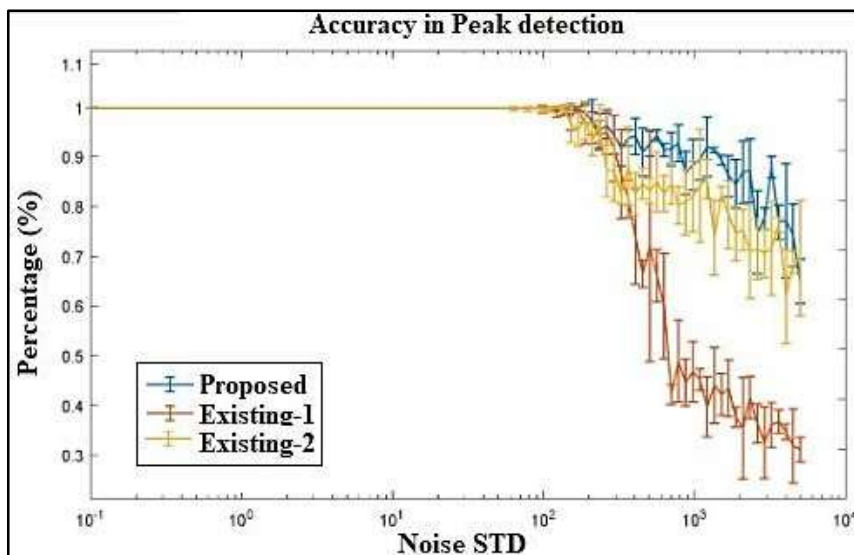


Figure 10. Accuracy obtained in detecting slope and peak of PPG signal

The arithmetic mean within the specificity and sensitivity of the delineation methods for various noise values are depicted in Figure-10 and Table-2

for the relevant fiducial locations: gradient and peak for each Spectro-Temporal and without Spectro-Temporal. Depending on the comparison of



the true positives (TPs), false positives (FPs), and false negatives (FNs) of the fiducial locations found by the delineation method on noise-corrupted data with the identification of a similar method on the actual data, the sensitivity, and specificity are calculated. As stated earlier, the efficiency of the proposed model is examined by the PPG signals, which are corrupted naturally by calculating the average pulse period. It is operated over an 8s window, and it is again compared with the average rhythm rate of the heart using a PPG signal. The comparison shows that the influence of spectro-temporal is obtained along with its outcomes with the phase of eliminating the PPG signal from the operating pipeline. The dataset is taken from [1] and used in this proposed approach. The error rates for predicting

the HR are shown in Table-3. It is preceded by using HR obtained from the PPG signal as the reference. Both the Spectro-Temporal and without Spectro-Temporal algorithms are applied. The results indicate that the Spectro-Temporal methodology minimizes the mean absolute error of HR. More precisely, the average mean absolute error is evaluated as 4.00 BPM for the corresponding Spectro-Temporal algorithm. The high-frequency resolution of Spectro-Temporal is highly beneficial to assessing ECG data where proper peaks can be selected. But the method is not efficient once the noise level gets increases. On the other hand, the smaller median error for Spectro-Temporal relative to without Spectro-Temporal is evidence that the improved accuracy reduces the variation of the outcome for various participants.

Table 2. Evaluation of Spectro-Temporal

Noise STD (σ)	Fiducial Location Identification Gmean (%)							
	Slope				Peak			
	Spectro-Temporal		Without Spectro-Temporal		Spectro-Temporal		Without Spectro-Temporal	
0	100	[0.00]	100	[0.00]	100	[0.00]	100	[0.00]
0.5	100	[0.00]	100	[0.00]	100	[0.00]	100	[0.00]
1	100	[0.00]	100	[0.00]	100	[0.00]	100	[0.00]
4	100	[0.00]	100	[0.00]	100	[0.00]	100	[0.00]
12	100	[0.00]	100	[0.00]	100	[0.00]	100	[0.00]
40	100	[0.00]	100	[0.00]	100	[0.00]	100	[0.00]
130	99.93	[0.08]	99.77	[-0.08]	98.43	[0.37]	98.36	[0.30]
450	93.47	[7.55]	90.65	[4.73]	89.68	[24.45]	83.33	[18.10]
1500	91.83	[25.71]	89.77	[23.65]	88.10	[47.18]	81.30	[40.38]
5000	77.70	[30.19]	77.89	[30.38]	63.70	[33.85]	68.38	[38.53]

Table3. Using heart rate error (BPM) from ECG (RR-intervals) and PPG (PP-intervals) on windows of 8 s, we may compare STCNN and fastSPARE

Subject	fastSPARE			STCNN		
	AE_{mean}	RE_{mean} (%)	AE_{median}	AE_{mean}	RE_{mean} (%)	AE_{median}
S1	3.80	4.37	1.38	6.26	7.41	1.50
S2	8.69	10.46	5.33	6.91	8.57	3.17
S3	3.29	3.13	2.30	3.22	3.18	2.39
S4	3.30	3.52	2.20	3.19	3.44	2.08
S5	4.31	4.94	1.78	6.69	7.61	2.21
S6	3.44	3.90	2.22	5.22	5.96	2.15
S7	2.62	2.77	1.54	2.58	2.72	1.62
S8	3.01	3.24	1.84	2.86	3.05	2.12
S9	2.90	2.90	2.02	2.58	2.59	1.75
S10	2.59	2.56	1.93	2.87	2.87	1.96
S11	3.55	3.83	1.80	3.24	3.49	1.97
S12	1.79	1.75	1.27	2.40	2.35	1.46
All	3.61	3.95	2.13	4.00	4.44	2.03

Table 4. Comparisons of Heart-Rate Error Detection



Subject	Spectro-Temporal (Mean of the Error)	Without Spectro-Temporal (Mean of the Error)
S1	3.80	6.26
S2	8.69	6.91
S3	3.29	3.22
S4	3.30	3.19
S5	4.31	6.69

Performance Evaluation

The overall experimental results estimate and verify some of the performance measures. The major measure used in this paper to evaluate the performance of the proposed spectral-temporal

analysis is F1-Score. Most of the learning algorithms are evaluated by calculating precision, recall, and accuracy, which are calculated from TP, TN, FP, and FN measures. This paper also does the same performance evaluation by calculating the F1-Score. Thus, the estimated F1-Score is compared with different convolutional methods for performance evaluation. The comparison of F1-Score is given in Table-1. It shows the performance of RF [1], Inception-V3 [2], ResNet-18 and 34 [3], and DenseNet [4], where these methods did not remove the sub-sampling layers initially. This paper also trained a CNN model with the same kind of structure explained in DenseNet [4].

Table 5. F1-Score Comparison

F1 _{overall}	Random Forest [1]	InceptionV3 [2]	ResNet18 [3]	ResNet34 [3]	DenseNet18 [4]	Proposed
CWT	74.91	76.41	78.57	78.70	78.82	96.63
BurgAR	73.22	76.45	76.41	76.30	77.58	87.76
FourierKS	75.99	77.48	78.05	77.99	79.50	85.24
OscKS	76.12	76.91	77.85	78.19	79.67	87.18

Regarding CWT, the proposed method explained here obtained 96.63% of F1-Score, which is higher than other algorithms. It also obtained better efficiency in terms of STFT, BurgAR, and Kalman techniques, where the proposed method gained more than 85% of F1-Score the other algorithms gained lesser than 80%. The proposed model is trained well using more input signals and manipulated using CWT. The learning rate of the proposed model is assigned as $1e^{-3}$ And the experiment is repeated for 50 epochs. The data normalization process is obtained using a batch size of 128. In addition to this, all the hyperparameters are tuned using hyperparameter-engineering over more number of layers.

The PPG signal leads the wearable field for heart rate evaluations due to its inconspicuous characteristics and very simple deployment in Internet-of-Things devices. However, its value extends beyond the straightforward recognition of the heart rhythm. In reality, this PPG signal can be employed to derive various indicators for multimodal systems for tracking the healthcare and wellbeing of individuals. PPG detectors are unfortunately susceptible to errors and disturbances anytime if the person moves slightly due to its basic operating mechanism. There are a variety of methods that many researchers have

proposed to tackle this problem. Many of them focus on predicting the HR of the individuals. Thus, we developed some advances in the existing system by adopting certain features depending on the characteristics of PPG signals. In this work, the reconstruction of the PPG signal waveform is created in the absence of accelerometers. The proposed method achieved better performance when compared with the other previous methods. In summary, the proposed innovative method reconstructs the heart's cardiac activity by representing it using a PPG signal waveform in three spectral elements. Moreover, this novel implementation is beneficial in compressing the PPG signal and examining it.

Conclusion

This paper has proposed a deep learning model for analyzing the spectro-temporal information of the PPG signals. Initially, it is motivated to remove the artifacts on the signals and reconstruct the full signal. It is shown that the spectro-temporal information of the PPG signal helps to identify various health problems, early symptoms, and noise data. It also helps to explore the physical properties of the signals. This paper decomposes and reconstructs the PPG signal for analysis efficiently. Many empirical values are shown for signals recorded under intensive physiological



movement and at ease and severely distorted by our simulated noise production to demonstrate the value of spectro-temporal data. In our research, spectro-temporal offered an enhancement for identifying several biomarkers of approximately 65%. As a result, spectro-temporal imaging offers good reconstruction, especially in signals completely damaged by distortions, enabling the recovery of several invaluable biomarkers.

References

- Zhilin Zhang, Zhouyue Pi, Benyuan Liu, "TROIKA: A General Framework for Heart Rate Monitoring Using Wrist-Type Photoplethysmographic Signals During Intensive Physical Exercise", *IEEE Trans. on Biomedical Engineering*, vol. 62, no. 2, pp. 522-531, February 2015.
- Zhilin Zhang, "Photoplethysmography-Based Heart Rate Monitoring in Physical Activities via Joint Sparse Spectrum Reconstruction", *IEEE Trans. on Biomedical Engineering*, vol. 62, no. 8, pp. 1902-1910, August 2015.
- Ding, X.; Yan, B.P.; Zhang, Y.T.; Liu, J.; Zhao, N.; Tsang, H.K. Pulse transit time based continuous cuffless blood pressure estimation: A new extension and a comprehensive evaluation. *Sci. Rep.* 2017, 7, 11554.
- Martínez, G.; Howard, N.; Abbott, D.; Lim, K.; Ward, R.; Elgendi, M. Can photoplethysmography replace arterial blood pressure in assessing blood pressure? *J. Clin. Med.* 2018, 7, 316.
- Hammerla, N.Y.; Halloran, S.; Ploetz, T. Deep, Convolutional, and Recurrent Models for Human Activity Recognition using Wearables. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, New York, NY, USA, 10 July 2016; pp. 1533-1540.
- Voisin, M.; Shen, Y.; Aliamiri, A.; Avati, A.; Hannun, A.; Ng, A. Ambulatory Atrial Fibrillation Monitoring Using Wearable Photoplethysmography with Deep Learning. *arXiv 2018*, arXiv:physics.med-ph/1811.07774.
- Biswas, D.; Simões-Capela, N.; Van Hoof, C.; Van Helleputte, N. Heart Rate Estimation From Wrist-Worn Photoplethysmography: A Review. *IEEE Sens.* 2019.
- Lee, B.; Han, J.; Baek, H.J.; Shin, J.H.; Park, K.S.; Yi, W.J. Improved Elimination of Motion Artifacts from a Photoplethysmographic Signal using a Kalman Smoother with Simultaneous Accelerometry. *Physiol. Meas.* 2010, 31, 1585-1603.
- Ram, M.R.; Madhav, VM; Krishna, E.H.; Komalla, N.R.; Reddy, K.A. A Novel Approach for Motion Artifact Reduction in PPG Signals Based on AS-LMS Adaptive Filter. *IEEE Trans. Instrum. Meas.* 2012, 61, 1445-1457.
- Kim, B.S.; Yoo, S.K. Motion Artifact Reduction in Photoplethysmography using Independent Component Analysis. *IEEE Trans. Biomed. Eng.* 2006, 53, 566-568.
- Sun, X.; Yang, P.; Li, Y.; Gao, Z.; Zhang, Y.T. Robust Heart Rate Beat Detection from Photoplethysmography Interlaced with Motion Artifacts based on Empirical Mode Decomposition. In *Proceedings of the International Conference on Biomedical and Health Informatics (BHI)*, Hong Kong, China, 5-7 January 2012; pp. 775-778.
- Zeng, M.; Gao, H.; Yu, T.; Mengshoel, O.J.; Langseth, H.; Lane, I.; Liu, X. Understanding and Improving Recurrent Networks for Human Activity Recognition by Continuous Attention. In *Proceedings of the International Symposium on Wearable Computers (ISWC)*, Singapore, 8-12 October 2018; 56-63.
- Hannink, J.; Kautz, T.; Pasluosta, C.F.; Gasmann, K.G.; Klucken, J.; Eskofier, B.M. Sensor-Based Gait Parameter Extraction With Deep Convolutional Neural Networks. *IEEE J. Biomed. Health Inform.* 2017, 21, 85-93.
- Jindal, V.; Birjandtalab, J.; Baran Pouyan, M.; Nourani, M. An Adaptive Deep Learning Approach for PPG-Based Identification. In *Proceedings of the IEEE EMBS International Conference on Engineering in Medicine and Biology*, Orlando, FL, USA, 16-20 August 2016; 6401-6404.
- Biswas, D.; Everson, L.; Liu, M.; Panwar, M.; Verhoef, B.E.; Patki, S.; Kim, C.H.; Acharyya, A.; Van Hoof, C.; Konijnenburg, M.; et al. CorNET: Deep Learning Framework for PPG-Based Heart Rate Estimation and Biometric Identification in Ambulant Environment. *IEEE Trans. Biomed. Circuits Syst.* 2019, 13, 282-291.
- Elgendi, M.; Norton, I.; Brearley, M.; Abbott, D.; Schuurmans, D. Systolic peak detection in acceleration photoplethysmograms measured from emergency responders in tropical conditions. *PLoS ONE* 2013, 8, 76585.
- Elgendi, M.; Norton, I.; Brearley, M.; Abbott, D.; Schuurmans, D. Detection of a and b waves in the acceleration photoplethysmogram. *Biomed. Eng. Online* 2014, 13, 139.
- Elgendi, M. Optimal signal quality index for photoplethysmogram signals. *Bioengineering* 2016, 3, 21.
- Elgendi, M.; Norton, I.; Brearley, M.; Fletcher, R.R.; Abbott, D.; Lovell, N.H.; Schuurmans, D. Towards investigating global warming impact on human health using derivatives of photoplethysmogram signals. *Int. J. Environ. Res. Public Heal.* 2015, 12, 12776-12791.
- Pan, J., & Tompkins, W.J. (1985). A real-time QRS detection algorithm. *IEEE Transactions on Biomedical Engineering BME*, 32(3), 230-236.
- Glorot, X., & Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the 13th International Conference on Artificial Intelligence and Statistics*, 9, 249-256.
- Huang, G., Liu, Z., van der Maaten, L., Weinberger, K.Q. (2017). Densely connected convolutional networks. In *2017 IEEE Conference on computer vision and pattern recognition (CVPR)*, 2261-2269.
- He, K., Zhang, X., Ren, S., Sun, J. (2016). Deep residual learning for image recognition. In *2016 IEEE Conference on computer vision and pattern recognition (CVPR)*, 770-778.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32.
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *The IEEE conference on computer vision and pattern recognition (CVPR)*.
- He, K., Zhang, X., Ren, S., Sun, J. (2016). Deep residual learning for image recognition. In *2016 IEEE Conference on computer vision and pattern recognition (CVPR)*, 770-778.
- Huang, G., Liu, Z., van der Maaten, L., Weinberger, K.Q. (2017). Densely connected convolutional networks. In *2017 IEEE Conference on computer vision and pattern recognition (CVPR)* (pp. 2261-2269).

