IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 7, 2022, accepted May 9, 2022, date of publication May 16, 2022, date of current version May 31, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3175436

Beat-Based PPG-ABP Cleaning Technique for Blood Pressure Estimation

MOSTAFA SALAH[1], OSAMA A. OMER[2], (Member, IEEE), LOAY HASSAN[3], MOHAMED RAGAB[2], AMMAR MOSTAFA HASSAN[4], AND AHMED ABDELREHEEM[5]

[1]Faculty of Engineering, Sohag University, Sohag 82524, Egypt
[2]Faculty of Engineering, Aswan University, Aswan 81528, Egypt
[3]Department of Information Systems, Universitat Rovira i Virgili, 43003 Tarragona, Spain
[4]Arab Academy for Science, Technology and Maritime Transport, South Valley Branch, Aswan 81511, Egypt
[5]Faculty of Computers and Information Technology, The National Egyptian Electronic Learning University, Aswan 81111, Egypt

Corresponding author: Osama A. Omer (omer.osama@aswu.edu.eg)

This work was supported in part by the Information Technology Industry Development Agency (ITIDA) under Grant ARP2019.R27.4.

ABSTRACT A growing attention is given to exploiting Photoplethysmography (PPG) signals in non-invasively measuring of many physiological vital signs. Many machine deep learning models were trained for predicting the continuous arterial blood pressure (ABP) or just the systolic and diastolic blood pressure (BP) values based on a public database. However, jointly cleaning the PPG-ABP dataset that is the most critical pre-processing step for training quality is still in need for more investigations. There is a considerable amount of anomaly data that has to be excluded before any training stage. This paper introduces a two-level joint PPG-ABP cleaning technique conducted in a signal level and per-beat level. Many quality metrics have been checked successively for excluding improper data. These metrics achieve a coarse cleaning step. Finally, principal component analysis (PCA) is exploited for fine cleaning the remaining data from the former stage. The cleaning efficiency is evaluated by measuring its impact on the deep-learning-based BP estimation models. The trained model based on our cleaned data shows performance enhancement in terms of prediction error and the correlation between the predicted and ground-truth BP. Segmented and cleaned PPG/ABP dataset will be publicly available in both signal level and beat level. Based on the simulation results, the proposed cleaning technique enhances standard deviation of the prediction error of systolic and diastolic blood pressure by 11.68 % and 10.81 %, respectively. Also, it enhances mean absolute error of the prediction of systolic and diastolic blood pressure by 14.79 % and 11.70 %, respectively.

INDEX TERMS Photoplethysmography, PPG-ABP dataset, data cleaning, PPG quality metrics, PPG, anomaly detection, PPG-ABP temporal/spectral alignment.

I. INTRODUCTION

Photoplethysmography (PPG) signal conveys information about changes in the volume of the blood in synchronous with heart pulsation mainly [1]. It can be measured simply using a contact fingerclip or wearable sensors. Also, it can be remotely estimated by a camera providing a remote-PPG (rPPG) signal. Wither being contact or remote PPG signal, it has been proved that PPG signal can be exploited for estimating important physiological vital signs such as blood pressure (BP) [2]–[9] oxygen saturation [10], and hemoglobin levels [11]. Even the multi-electrode electrocardiogram (ECG) signal, can be inferred via PPG signals [12]. These techniques guarantee the intended simplicity of non-clinical monitoring of health status. Thanks to the public availability of biomedical datasets [13]–[15] that enable training deep learning models [5] for inferring the relation between these data and the corresponding labeled physiological parameters. However, these huge datasets need to be cleaned effectively before any further deployment. It is collected and archived without paying attention to its quality. Hence, many techniques were introduced for data cleaning and abnormality/artifact detection [16]–[22].

By regarding our main interest in PPG-ABP dataset for BP estimation, there are still many opportunities for further data cleaning. Other than BP estimation based on pulse arrival time (PAT) techniques [23]–[26] where the information resides in the phase shift between PPG and ECG signals,

The associate editor coordinating the review of this manuscript and approving it for publication was Kostas Kolomvatsos.

BP estimation from PPG only [4], [27] imposes higher challenges in model fitting. All information is learned from the PPG morphological shape only. The quality of the training data has a critical impact on the performance of the deep-learning-based training models. Defected PPG training beats lead to unreliable modeling of BP. Model accuracy is highly dependent on training data quality.

In this paper, we are introducing a joint PPG-ABP cleaning technique in two successive levels; signal level and per-beat level. The first level is carried on signal level. The spectral alignment between PPG and ABP signals is applied. Then, both PPG and ABP sets are segmented into beats. Some quality measures are imposed based on physiological thresholds. Each quality metric excludes apart from the defected beats successively. The defective PPG/ABP beat will be dropped out along with its corresponding beat PPG/ABP where the ABP beat represents the labeling of the corresponding PPG beat. To that end, coarse data cleaning is performed. Then, a simple principal component analysis (PCA) -based fine cleaning is applied. The proposed cleaning approach is evaluated by measuring the standard deviation and the mean of error of the estimated BP through some of the cleaning stages.

The main contributions in this paper can be summarized in:

- Introducing a two-level cleaning technique for PPG and ABP signals.
- Double cleaning strategy: Physiological limits are applied on both PPG/ABP signals/beats. If one signal/beat is defective, both signals/beats will be excluded for providing both trusted data and its label as well.
- Temporal/spectral alignment: Data consistency is validated by checking the minimal time/spectral correlation between PPG and ABP signals/beats.
- Introducing fine cleaning based on principal component analysis (PCA) representation.
- Joint PPG/ABP beat-by-beat cleaned dataset is made publically available.

II. RELATED WORK

One of the main problems obstructing full exploitation of the available biomedical databases resides in archiving these data without any concern about its quality. The percentage of the defective beats is very large that is making any training based on these data unreliable. A distinct trial for assessing the data along its collection is conducted in china over 219 subjects only [28]. Other than that trial, there are many techniques for providing quality measures or just detecting abnormal signal intervals.

The signal abnormality index (SAI) algorithm [22] detects abnormality intervals in the ABP signal. It applies successive series of abnormality metrics based on the limits of living physiological parameters such as the extreme values of systolic and diastolic BP, the maximum difference between systolic and diastolic values and the extreme limits of heart rate. Detection of valid ABP beats was enhanced by combining the former SAI to signal-to-noise (SNR) measure. SNR is estimated as the percentage of the representation of the ABP beat in signal space compared to the noisy complementing space [29]. Signal space is spanned by the aid of singular-value decomposition (SVD) of 567 visually validated beats. In [30], correlation to updated template beat is employed for per-beat artifact detection. In [6], besides physiological criteria, template matching is employed for data cleaning. In [18], signal quality assessment for ABP and ECG signal is discussed based on physiological and morphological limits. The authors main aim was to reduce the false alarm rate in the monitor of the intensive care unit (ICU). In [17], Elgendi studied many metrics for measuring the PPG signal quality index (SQI). He has concluded that the skewness metric outperforms other metrics in determining the signal quality. Also, combined metrics improve the overall performance.

Neural network-based PPG classification is presented in [20]. It provides signal labeling based on the difference between the estimated heart rate from PPG and ECG signals. Ten beats are the upper bound difference that can be allowed between the two signals for good labeling. However, this simple labeling ignores the minimum level of morphological quality completely. Both Machine learning and deep-learning techniques [16], [17], [21], [31], [32] are employed for data cleaning or classifying signals as acceptable or anomalous. However, most of these approaches rely on manual data labeling based on experts' annotation. Hence, it was performed on a limited number of signals. Also, the assessment relies on the annotators' view. Pure manual classification is presented in [19]. It is based on gathering views of 18 international researchers for determining the required minimum recording length and the minimum number of undistorted pulses.

PPG quality may be classified based on diagnosis possibility into; excellent (salient systolic and diastolic), acceptable (not salient systolic and diastolic along with determined HR), and unfit (noisy signal) [17]. However, rather than just detecting defective beats/signal intervals, we are interested in extracting trusted clean data with clear features for training deep learning models effectively.

III. MATERIALS AND METHODS

This paper provides a joint PPG-ABP cleaning technique. It presents a consecutive two-level signal/per-beat cleaning for both PPG and the corresponding ABP signals where the ABP signal conveys the BP label of the corresponding PPG signal. So, any defective PPG/ABP beats will lead to exclusion of both of PPG and ABP signals.

A. DATA SET

Thanks to the Physionet's MIMIC II dataset (Multi-parameter Intelligent Monitoring in Intensive Care) [13] that provides joint PPG-ABP data needed for feeding the deep learning models. A compiled version of that dataset[1] is introduced in [14] where it has a better presentation. It is assumed that this

[1]Available at https://archive.ics.uci.edu/ml/datasets/Cuff-Less+Blood+Pressure+Estimation. Accessed at may 2021.

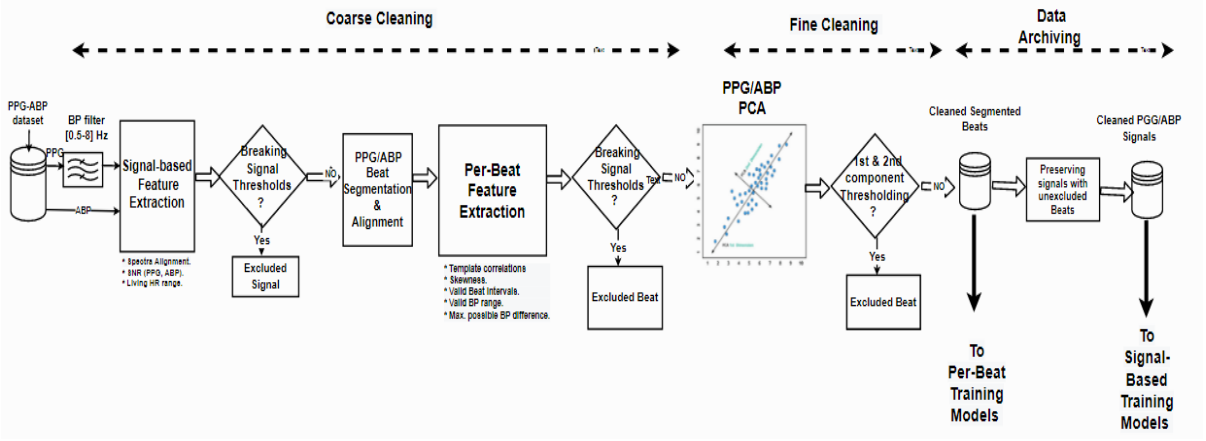


FIGURE 1. Proposed PPG-ABPCleaning technique.

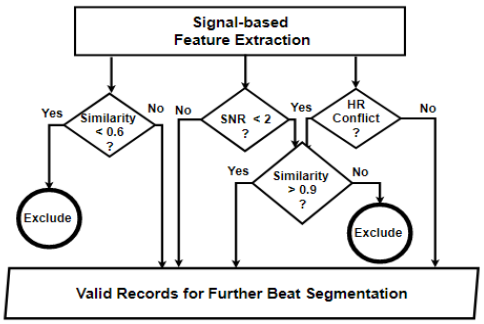


FIGURE 2. Logic of signal-level cleaning.

dataset was cleaned. However, by inspecting that dataset, it is shown that it still has a considerable amount of defective PPG and ABP signals. That dataset represents our main material that will be utilized for providing a jointly cleaned PPG-ABP dataset for feeding deep learning BP estimation models.

That dataset contains 12,000 records of different lengths. Each record includes ABP (invasive arterial blood pressure (mmHg)), PPG (photoplethysmograph from fingertip), and ECG (electrocardiogram from channel II) signals sampled at $F_s = 125$ samples/sec. However, we are interested only in the PPG signals and their corresponding ABP signals as a label. For proper handling and filtering, records are segmented into sections of 1024 samples length. So, we have 30,660 records.

B. PPG/ABP PRE-PROCESSING AND ALIGNMENT

Only PPG signals can be pre-processed by enhancing technique such as band pass filtering in the band [0.5-8] Hz as long as its morphological shape does not alter. On the other hand, the ABP signal cannot be touched because any trial for enhancing its quality leads to changes in its magnitude that represents the BP value. Also, in the beat segmentation stage, PPG beats will be normalized while the corresponding ABP beats remain untouched. Heavily distorted ABP signals or beats have to be excluded.

It worth emphasizing that there is some phase difference between the two signals corresponding to the difference in the sensing locations. Therefore, PPG and ABP signals alignment is needed for measuring similarity between the two signals and for segmentation purposes. However, this phase shift does not exceed the beat interval. So, to maintain correct beat correspondence, the ABP signal is shifted to get the maximum cross-correlation with the PPG signal as follows.

$$x_{cor}(k) = \sum_{n=1}^{N} ABP\,[n]\,PPG\,[n + \mathrm{k}], \tag{1}$$

For the maximum shift $0 \leq \mathrm{k} \leq \frac{\mathrm{F_s}}{2}$, where N is the beat length. It is predicted that the maximum shift between the two signals does not exceed half second.

C. MOTIVATIONS

PPG signal quality can be measured based on many signal quality indices [17] such as spectral signal-to-noise ratio (SNR), skewness, and Kurtosis. Then, data cleaning may be converted into designing proper thresholds for these metrics. Signals that are breaking these thresholds can be excluded. However, strict cleaning based on the signal level only (full-length PPG/ABP signal) seems improper for the following reasons:

1) Long recording of Physiological Data is highly susceptible to interference or distortion in some intervals due to sensor/patient movement or misconnection.
2) Intervals of defective data have a negative impact on the features related to the whole signal.
3) Excluding the whole signal that includes some defective beats leads to huge wastage in data resources.
4) For instance, a strict signal-based data cleaning algorithm that was applied in [7] on MIMIC III database, reduces the number of validated patient records from over 30000 to about 510 records only.
5) Under remote (camera-based) rPPG detection, the resulting SNR is too low to provide a long signal with acceptable features. It may be sufficient to predict BP from a few beats extracted with clear features.

Therefore, the main goal in this paper is to present a cleaning technique that can be used dataset cleaning for training as well as for prediction by testing the estimated rPPG beats.

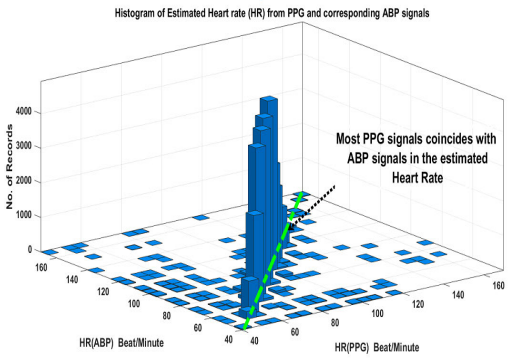


FIGURE 3. Histogram of estimated heart rate (HR) from PPG and corresponding ABP signals. Most signals agree in the estimated HR. There are 544 excluded records due to HR estimation conflict between PPG and ABP signals.

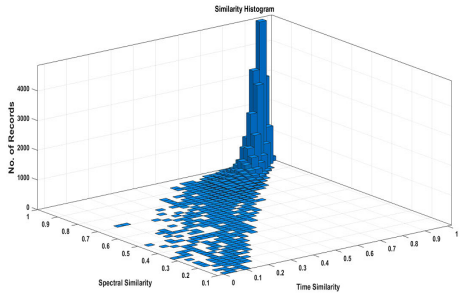


FIGURE 4. Histogram of time/spectral similarity between PPG and ABP signals. Note: Time similarity is computed at both signal alignment. Also, Spectral similarity is computed based on magnitude spectra.

IV. PROPOSED METHODOLOGY

In this paper, we follow a successive two-level cleaning technique. Signal level cleaning is based on mild thresholding excludes signals with salient abnormality. Per-beat cleaning is applied by thresholding many per-beat features. In both levels, there is no unique metric that can provide perfect cleaning. Defected signals/beats will be identified based on the combination of some successive metrics. It worth mentioning that the cleaning is performed on both PPG and ABP signals. In either case, physiological limits are also applied on both signals in the signal and beat levels.PPG/ABP signals and beats less than assumed thresholds will be dropped out along with their corresponding ABP/PPG records. To this end, coarse cleaning has been achieved. The final fine cleaning stage relay on PCA thresholding. Figure 1 shows the proposed cleaning sequence entirely. The details of the proposed cleaning technique are explained in the following subsections.

A. SIGNAL LEVEL CLEANING

In this stage, both of PPG and ABP signals are inspected without any segmentation. The main idea resides in viewing both signals in the frequency domain where both signals have to exhibit similar spectral construction as follows. The exclusion logic for single-level cleaning is presented in Fig. 2.

1) SPECTRA ALIGNMENT PRINCIPAL

Both PPG and ABP signals are arising from the same pulsating source that is the heart. So, PPG and ABP signals can be considered as quasi-periodic signals with identical fundamental frequencies. Hence, the estimated heart rate from the PPG signal (HR_{ppg}) has to coincide with that estimated from the ABP signal (HR_{abp}). To account for spectral resolution and possible errors in the spectral estimation process, a tolerance of ten beats can be set as a maximum deviation between the two heart rates. So, records can be excluded for the condition stated in Eq.2, which refers to the HR estimation conflict.

$$\left|HR_{ppg} - HR_{abp}\right| > 10 \tag{2}$$

Figure 3 shows the estimated HR histogram where most of the PPG/ABP signals are similar in the estimated HR. A small number of records exhibit HR conflict. These records will be tested by similarity check before exclusion decision.

2) HR OUT OF POSSIBLE RANGE

Physiological limits impose upper and lower bounds on the HR ranges. Often, the estimated fundamental frequency (F_0) determines the estimated heart rate. Hence, the constraints on HR imply corresponding constraints on the spectral shape. So, records can be excluded under the following conditions

$$HR \leq 40\ beat/minute \equiv F_0 \leq 0.667\ Hz \tag{3}$$
$$HR \geq 180\ beat/minute \equiv F_0 \geq 3\ Hz \tag{4}$$

3) SNR THRESHOLDING

Traditionally, the signal-to-noise ratio (SNR) represents a common metric for measuring signal quality. Thanks to the semi-periodicity of PPG/ABP signals, most of the signal power is concentrated around the fundamental frequency and its harmonics with a very narrow bandwidth; about 0.2 Hz. (5), as shown at the bottom of the page, where $\hat{P}_f$ represents spectral power density measured over the cardiac band and $\Omega = [0.5\text{–}8]$ Hz. The impact of any defective beat is reflected directly on the SNR level. As mentioned before, SNR thresholding is applied on PPG/ABP signals for excluding time intervals over which there is salient distortion. However, a light SNR threshold is followed for preserving undistorted beats from signals that have one or two defected beats only. Originally, the ABP signal cannot be filtered, the whole spectra band $\Omega = [0 - \frac{F_s}{2}]$ Hz contributes to the

$$Q_{\text{SNR}} = \frac{\int_{f_0-\Delta f}^{f_0+\Delta f} \hat{P}_f \mathrm{df} + \int_{2f_0-\Delta f}^{2f_0+\Delta f} \hat{P}_f \mathrm{df} + \int_{3f_0-\Delta f}^{3f_0+\Delta f} \hat{P}_f \mathrm{df}}{\int_{\Omega} \hat{P}_f \mathrm{df} - \left(\int_{f_0-\Delta f}^{f_0+\Delta f} \hat{P}_f \mathrm{df} + \int_{2f_0-2\Delta f}^{2f_0+2\Delta f} \hat{P}_f \mathrm{df} + \int_{3f_0-\Delta f}^{3f_0+\Delta f} \hat{P}_f \mathrm{df}\right)} \tag{5}$$

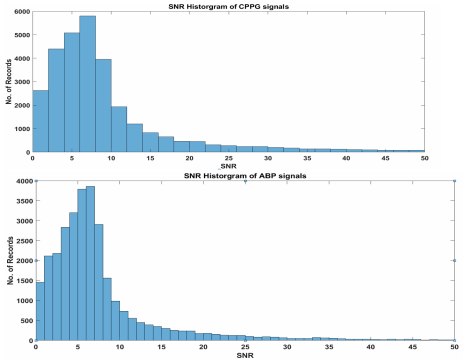


FIGURE 5. SNR histogram for PPG and ABP signals.

noise level except the components around the fundamental and the first two harmonics. The offset DC of the ABP signals will be ignored in computing the noise level because it is necessary for that signal. However, it does not contribute to the pulsating activates which represents the signal of interest. Signals with very low SNRs will be dropped because there is a low probability of extracting good beats from these signals.

4) PROBLEMS IN DETECTING HR

The exact estimation of HR has a critical impact on the overall assessment of signal quality. Wrong HR estimation fires false alarm of HR conflict and leads to underestimation of the SNR of these signals. Consequently, these signals may be excluded and lost. Other than remote (camera-based) signal detection, we assumed that we have signals extracted from contact sensors where the HR is salient enough to be detected without any special processing efforts. Simply, we rely on Fourier spectral estimation for detecting the fundamental frequency (F_0) bin as the largest magnitude spectra component in the signal representation and the harmonics are located exactly at an integer multiple of the fundamental frequency. Often, the spectrum exhibits a sparse structure with high power concentration around the fundamental frequency and its harmonics. Unfortunately, under these simple assumptions, we encountered by false alarm of 1490 records with HR out-of-human band, and 2538 records with HR conflicts between the estimated HR from PPG and the corresponding ABP signals.

By inspecting these records, we have encountered other cases that break these naive assumptions:

1) The 1st or 2nd harmonic may have larger magnitudes than the fundamental frequency.
2) Fundamental component and the assumed harmonics may be observed out of perfect harmony that is not an exact integer multiple of F_0.
3) Continuous (non-sparse) spectrum that arises from heart rate variability.

So, for accounting for these observations, the algorithm starts by detecting the largest spectral bin in the possible range [0.665- 3] Hz. Then, the spectral magnitude of possibly frequency bins for harmonic construction is checked for being in harmony or not. By these adaptations, the number of HR conflicts is reduced to 544 records without detecting any HR out-of-human band. Any detected case of HR conflict includes signal distortion either in PPG, ABP, or both. The detected conflict in HR is a sufficient indicator for dropping that record completely.

5) TIME/SPECTRAL SELF-SIMILARITY

As reported in [4], [33], there are strong morphological correlations between PPG and corresponding ABP signals in both time and frequency domains. It worth mentioning that both signals have to be aligned (shifted in phase) before measuring time similarity.

Time similarity is determined as follows,

$$Time\_similarity = \frac{\sum (PPG - m_{PPG})\,(ABP - m_{ABP})}{\sqrt{\sum (PPG - m_{ABP})^2 \sum (ABP - m_{ABP})^2}}$$

where m_{PPG} and m_{ABP} are the mean values of the PPG and ABP signals, respectively.

Also, spectral correlation is performed on the magnitude spectra of both signals. Figure 4 shows a joint correlation histogram where most of the signals exhibit a high similarity larger than 0.8. Similarity measures enhance cleaning performance in two ways; 1) Excluding Signals with low correlation coefficients, and 2) Re-assessment of signals violating previous metrics (SNR and HR conflict) for saving incorrectly classified signals from exclusion.

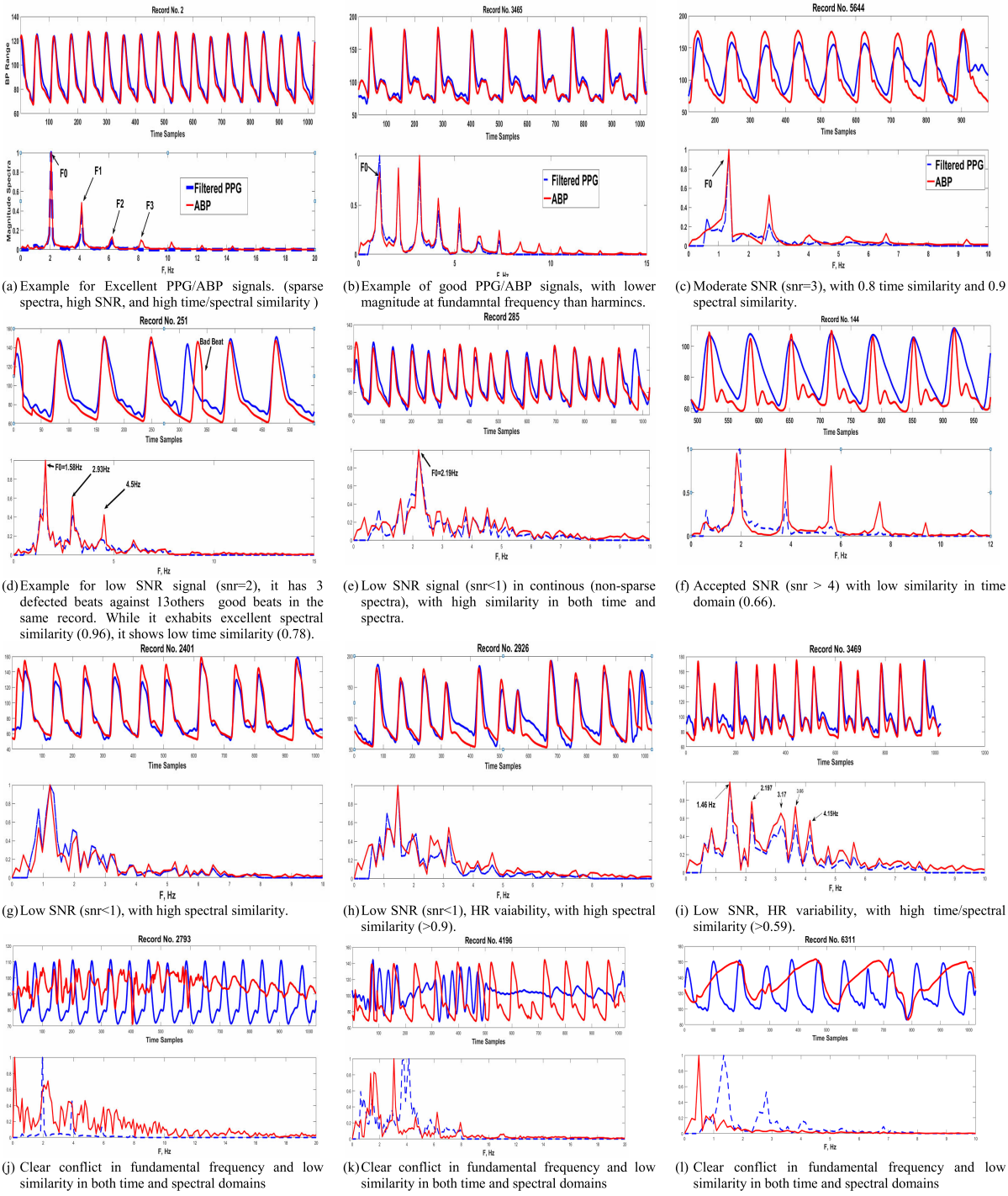
So, we assume that if the signal fails in satisfying the SNR threshold while exhibiting extremely high similarities in both time and spectral domains, that signal should not be excluded. That is because it is not highly probable to have the two signals (that are measured from independent sensors) jointly distorted and producing similar pulsation. Moreover, one or two bad beats may degrade the overall SNR of the record. These few beats will be dropped in the next stage. Also, SNR criteria fail in estimating signal quality effectively under heart rate variability. It worth mentioning that the similarity criterion is not a necessary condition for accepting the signal as a valid signal however, it is a sufficient indication for inspecting it carefully before taking a exclusion decision.

B. PER-BEAT CLEANING LEVEL

Per-beat cleaning is more effective than signal level cleaning because only defective beats will be dropped out while good beats can be preserved from the same signal. The PPG/ABP signals will be segmented into beats where each ABP beat represents the label for the corresponding PPG beat. Then, some cleaning metrics will be applied for excluding defective beats.

1) PPG/ABP SEGMENTATION

Segmentation is an effective task in the cleaning stage because improperly segmented beats will be misclassified as a bad beat. There are many efforts in the PPG beat segmentation area for providing accurate segmentation in presence of artifacts and distortion [34]–[36]. However, in this work,



(a) Example for Excellent PPG/ABP signals. (sparse spectra, high SNR, and high time/spectral similarity)

(b) Example of good PPG/ABP signals, with lower magnitude at fundamntal frequency than harmincs.

(c) Moderate SNR (snr=3), with 0.8 time similarity and 0.9 spectral similarity.

(d) Example for low SNR signal (snr=2), it has 3 defected beats against 13others good beats in the same record. While it exhabits excellent spectral similarity (0.96), it shows low time similarity (0.78).

(e) Low SNR signal (snr<1) in continous (non-sparse spectra), with high similarity in both time and spectra.

(f) Accepted SNR (snr > 4) with low similarity in time domain (0.66).

(g) Low SNR (snr<1), with high spectral similarity.

(h) Low SNR (snr<1), HR vaiability, with high spectral similarity (>0.9).

(i) Low SNR, HR variability, with high time/spectral similarity (>0.59).

(j) Clear conflict in fundamental frequency and low similarity in both time and spectral domains

(k) Clear conflict in fundamental frequency and low similarity in both time and spectral domains

(l) Clear conflict in fundamental frequency and low similarity in both time and spectral domains

FIGURE 6. Examples of classified records in signal-level cleaning.

we follow the simple algorithm described in [37] for determining beats limits as local minima points.

First, PPG/ABP signals are aligned (as already explained in the pre-processing stage) for preserving accurate beat correspondence. Then, beat marks are estimated for one signal (PPG) then determining the beat limits for the other signal (ABP) by simple fine adaptation around the previously determined marks. Figure 7 demonstrates the alignment, beat marking, and beat segmentation steps.

Segmentation failure may occur due to beat artifacts, beat class, or algorithm weakness as shown in Fig.8 where improper segmentation leads to strange PPG beats as in Fig.8.b. According to the observation study [38], the PPG waveform is classified according to the amplitude and the position of the diacritic notch. Among these classes, vasodilatation case exhibits low or negative notch, hence, this class usually experiences poor segmentation. Anyway, the performance of the segmentation algorithm is out of the scope of our

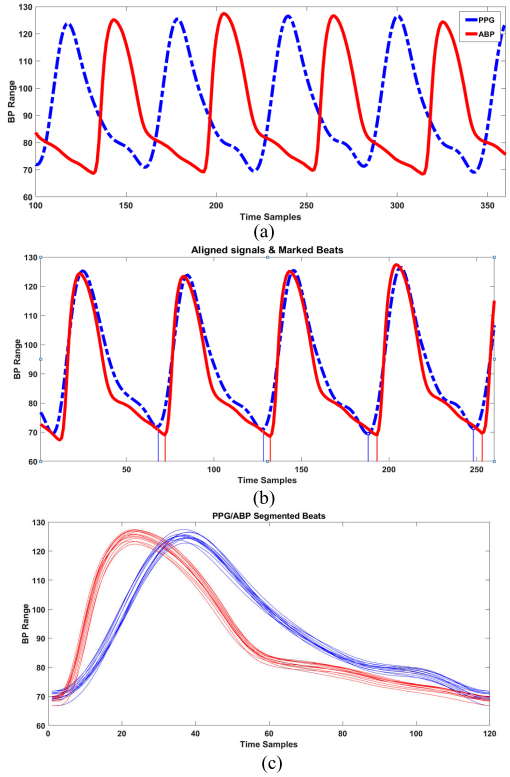


FIGURE 7. Signal segmentation; a) Normal PPG/ABP signal; b) Aligned PPG/ABP signals with marked beats; c) PPG/ABP segmented beats along time normalization.

paper. Our main focus is to exclude badly segmented beats in the following stages irrespective to the error source.

2) SUCCESSIVE BEAT SIMILARITY

Regular PPG/ABP signals exhibit semi-regular pulsation with minimal variation between successive beats. Abrupt morphological change in successive beats may indicate sensor movement or unstable connection to the body. Hence, it has to be excluded. Usually, the first and last beats are dropped because dividing the signal into a fixed length of 1024 samples often induces incomplete beats. Then, beats exhibiting low similarity with others in the same section will be dropped.

3) PPG/ABP BEAT SIMILARITY

As reported [4], [39], there is a high coherence between PPG and ABP beats. Based on this coherence, deep learning models [3], [40] are induced for learning PPG/APB signal translation. Then, systolic (sysBP) and diastolic (DiaBP) can be estimated from the translated signal for each beat couple (ppg_i, abp_i) as follows

$$Sys_{BP}\left(Beat_i\right) = Max\left(abp_i\right)$$
$$Dia_{BP}\left(Beat_i\right) = Min\left(abp_i\right)$$

where the systolic is the maximum BP and the diastolic is the minimum in the abp beat. Note; we use PPG/ABP for denoting signals (1024 samples), while ppg/abp stands for beats resized into 120 samples long.

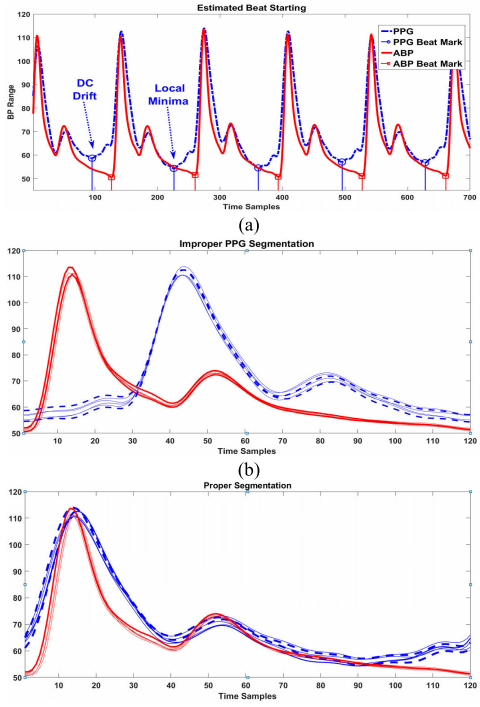


FIGURE 8. Example of the critical impact of segmentation, a) PPG/ABP signal, b) Improperly segmented PPG beats, c) Proper segmented PPG beats.

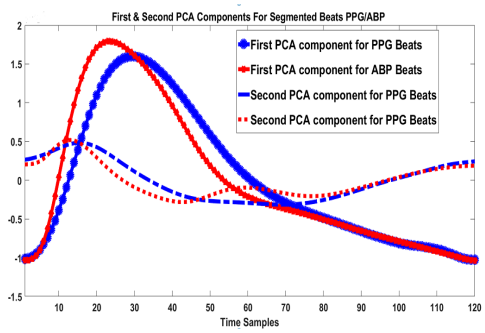


FIGURE 9. The first and second PCA components for PPG/ABP segmented beats.

Hence, a very low correlation between corresponding beats of the PPG and ABP signals indicates bad beats that should be excluded.

a: TIME NORMALIZATION

Beat length varies from one subject to another. Also, it may vary for the same subject. Hence, segmented beats have different lengths that can't be fed into a fixed-length training models. Also, we need to embed the beat interval as an essential feature that cannot be neglected. So, each beat is allowed to have a different length in the cardiac range (0.33 – 1.5 sec). Time normalization can be performed by unifying the beat length to 120 sample/beat that corresponds to the maximum possible beat interval (1.5 sec). Consequently, the beat intervals are included as an extra feature so the length of the input data for each beat that is used for training, validation and testing is 121 samples.

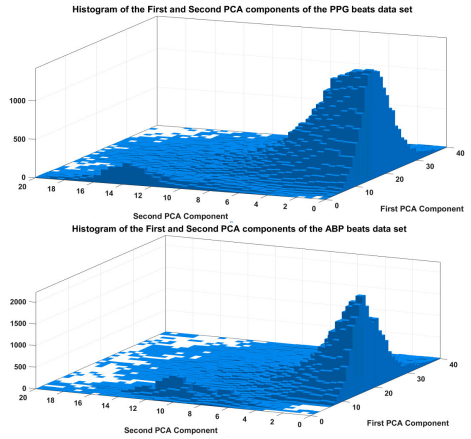


FIGURE 10. Biplot histogram of the first and second PCA components for PPG/ABP segmented beats.

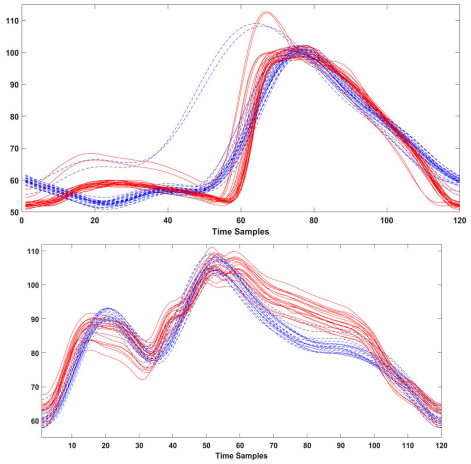


FIGURE 11. Examples of detected beat artifacts by PCA cleaning.

b: AMPLITUDE NORMALIZATION

Only the segmented PPG beats can be normalized to the amplitude [0-1] range.

$$ppg_{Normalized} = \frac{ppg - \min ppg}{\max ppg - \min ppg} \tag{6}$$

While the ABP beats are kept without any modifications.

C. PCA-BASED FINE CLEANING

To this end, we have PPG/ABP segmented beats dataset. Many metrics are imposed for excluding abnormal beats. It is supposed that these beats have an accepted level of similarity between, PPG and ABP beats, and also between the PPG beats from the same record. However, there is no guarantee that all these beats are valid and belong to the same natural generating phenomena. That dataset will be employed as a training set for deep-learning models. So, the BP predicting performance relies mainly on the quality of the training dataset. Hence, more cleaned data with reduced anomaly percentage leads to better-trained models. In the first part of the work, we have applied a combination of metrics for excluding bad beats/signals. Now, the remaining set will be employed for cleaning itself.

Principal component analysis (PCA) [41] can be used for representing high dimensional data into smaller set orthogonal components. Often, the first and second components convey about 70% of data variability/variance, whereas the first component has a higher contribution in the signal representation than second component. Based on this point, simple outlier detection can be performed through a biplot of these components as in [42]. Figure 9 shows the first two PCA components for both PPG/ABP sets.

Figure 10 demonstrates the biplot of first and second PCA components for PPG and ABP sets. By inspecting that biplot, we note that almost all the data is concentrated in certain distribution that should represent good beats. It seems an unsupervised classification strategy. This step excludes about 6% of the segmented beats. An example of the excluded PPG/ABP beats is shown in Fig.11. Some of the excluded beats exhibit systolic lower than diastolic BP as in Fig. 11.

TABLE 1. Standard deviation and mean error of systolic/diastolic BP prediction.

Cleaning Stage	SBP		DBP	
	STD	Mean	STD	Mean
All Segmented Beats	22.588	17.58	11.956	8.8129
After Signal & Beat cleaning	22.008	18.17	11.419	9.733
After PCA clean	19.95	14.98	10.663	7.782

D. QUALITY MEASURE

The impact of the PPG/ABP cleaned dataset is not limited to feeding the training models for BP learning from corresponding PPG. That cleaned set can be employed for assessing any PPG signal/beat quality. Hence, any PPG signal (extracted from contact PPG or remotely sensed from a camera) can be validated as an artifact-free signal/beat. The subspace spanned by the first PCA components can be employed for measuring beat quality. Normal beat shapes belong to that subspace. The distance between the beat and the subspace represents how far that beat is from the normal beat shape.

V. CLEANING EVALUATION

The cleaning efficiency represents the percentage of the normal beats to the whole number of beats in the dataset. However, it cannot be measured directly. But, it can be evaluated by measuring its impact on the accuracy of the deep-learning-based BP estimation models.

The long short-term memory (LSTM) model [43] is employed for training our dataset. As mentioned before, all segmented beats are resized into 120 sample lengths while the beat interval is embedded in the sample #121 in the input data. The model will be trained for systolic prediction from segmented PPG beats.

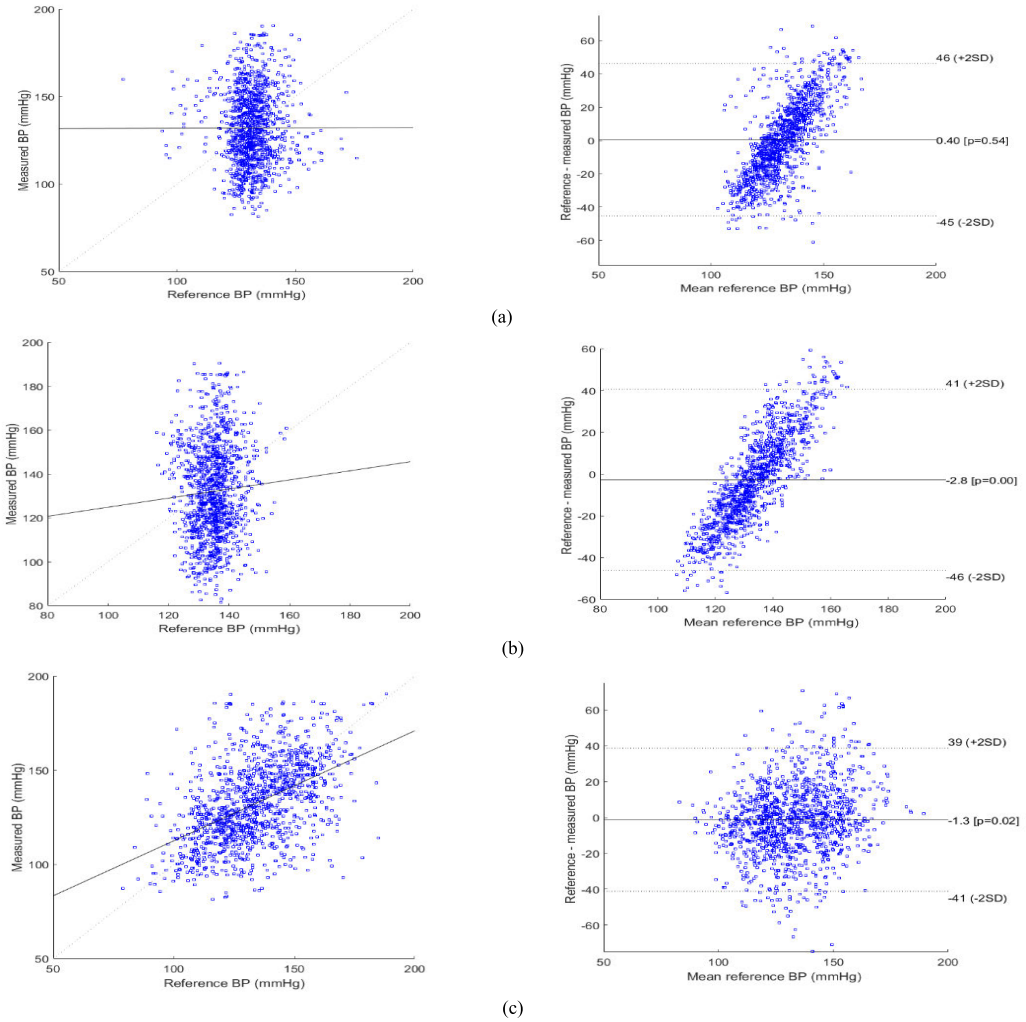


FIGURE 12. Bland–Altman plot for cleaning evaluation at different cleaning stages. a) All segmented beats without cleaning, b) After signal/beat cleaning, c) after PCA cleaning.

The LSTM model is trained independently by the extracted/segmented PPG beats datasets in three different stages:

1) All segmented beats before cleaning.
2) After signal and beat cleaning stages (coarse cleaning).
3) After PCA fine cleaning (final stage).

Table 1 summarizes the standard deviations and mean errors encountered in the systolic BP prediction of segmented beats for three cleaning states of the dataset. Based on these results, it can be shown that the proposed cleaning technique enhances standard deviation of the prediction error of systolic and diastolic blood pressure by 11.68 %, 10.81 %, respectively. Also, it enhances mean absolute error of the prediction of systolic and diastolic blood pressure by 14.79 %, 11.70 %, respectively.

Figure 12 demonstrates the performance enhancement in terms of the Bland-Altman plot and correlation plot. As it is shown in Fig. 12c, the two-levels cleaning leads to more correlation between the predicted BP and the ground truth values. Also, the error in the dtandard range achieved by the proposed cleaning technique is lower than that achieved by uncleaned data. In addition, the error achieved by the proposed cleaning technique is less spreading compared to that of the uncleaned data. Moreover, to achieve a standard deviation of 2, the range of error achived by the proposed cleaning technique is $-41\sim39$ while using uncleaned data achieved error range $-45\sim46$. It worth to mention that the dataset is reduced to 175000 beats after the proposed cleaning technique compared to 309000 beats uncleaned data. From these results, it can be demonstrated that uncleaned data can make illusion for the deep learning network even with using huge data set. On the other hand, using less number of datasets can lead to better learning for the deep learning networks if properly cleaned.

VI. DATA AVAILABILITY

Our cleaned and pre-processed data are available at: https://cibpm.com/

VII. CONCLUSION

This paper introduces a joint PPG-ABP cleaning strategy. The key idea resides in making use of the fact that both

signals are arising from the same cardiac source. Hence, time/spectral similarity validates the signal as a normal signal and allows simple exclusion of defective signals/beats. The cleaning is performed on two levels; signal level and beat level. Signal level cleaning alone is not enough. One or two defected beats degrades the signal-to-noise ratio of the whole signal. So, in the beat level cleaning, bad beats are dropped only while good beats are preserved. There is no unique metric that can identify anomaly signals/beats. A combination of successive metrics is applied for coarse cleaning. Fine cleaning is performed based on the data itself on beat-level. It is supposed that a coarse cleaned dataset has a small percentage of anomalies while the major part of the dataset is assumed valid data that forms the PCA representation. Anomaly detection is adapted as unsupervised classification based on the biplot of the first two PCA components. Cleaning evaluation is performed by noting the enhancement in BP learning according to the cleaning degree. The power of cleaned data can be extended to provide a quality quantifying technique of observed PPG signals/beats where low-quality signals/beats should not be employed for further BP estimation. The cleaned dataset is made available at both signal and beat levels.

ACKNOWLEDGMENT

Authors would like to thank Advanced Computer Systems (ACS) Company for their continuous support. Also, they would like to thank Eng. Abdel-Rahan Abol-Kassem, the technician in their laboratory, for his continuous technical support.

REFERENCES

[1] M. Elgendi, "On the analysis of fingertip photoplethysmogram signals," *Current Cardiol. Rev.*, vol. 8, no. 1, pp. 14–25, Feb. 2012.

[2] Y.-C. Hsu, Y.-H. Li, C.-C. Chang, and L. N. Harfiya, "Generalized deep neural network model for cuffless blood pressure estimation with photoplethysmogram signal only," *Sensors*, vol. 20, no. 19, p. 5668, Oct. 2020.

[3] N. Ibtehaz and M. S. Rahman, "PPG2ABP: Translating photoplethysmogram (PPG) signals to arterial blood pressure (ABP) waveforms using fully convolutional neural networks," 2020, *arXiv:2005.01669*.

[4] G. Martínez, N. Howard, D. Abbott, K. Lim, R. Ward, and M. Elgendi, "Can photoplethysmography replace arterial blood pressure in the assessment of blood pressure?" *J. Clin. Med.*, vol. 7, no. 10, p. 316, Sep. 2018.

[5] F. Schrumpf, P. Frenzel, C. Aust, G. Osterhoff, and M. Fuchs, "Assessment of deep learning based blood pressure prediction from PPG and rPPG signals," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2021, pp. 3820–3830.

[6] G. Slapničar, M. Luštrek, and M. Marinko, "Continuous blood pressure estimation from PPG signal," *Informatica*, vol. 42, p. 1–15, Mar. 2018.

[7] G. Slapničar, N. Mlakar, and M. Luštrek, "Blood pressure estimation from photoplethysmogram using a spectro-temporal deep neural network," *Sensors*, vol. 19, no. 15, p. 3420, Aug. 2019.

[8] X. Xing and M. Sun, "Optical blood pressure estimation with photoplethysmography and FFT-based neural networks," *Biomed. Opt. Exp.*, vol. 7, pp. 3007–3020, Aug. 2016.

[9] W.-R. Yan, R.-C. Peng, Y.-T. Zhang, and D. Ho, "Cuffless continuous blood pressure estimation from pulse morphology of photoplethysmograms," *IEEE Access*, vol. 7, pp. 141970–141977, 2019.

[10] P. M. Mohan, A. A. Nisha, V. Nagarajan, and E. S. J. Jothi, "Measurement of arterial oxygen saturation (SpO2) using PPG optical sensor," in *Proc. Int. Conf. Commun. Signal Process. (ICCSP)*, Apr. 2016, pp. 1136–1140.

[11] A. R. Kavsaoğlu, K. Polat, and M. Hariharan, "Non-invasive prediction of hemoglobin level using machine learning techniques with the PPG signal's characteristics features," *Appl. Soft Comput.*, vol. 37, pp. 983–991, Dec. 2015.

[12] Q. Zhu, X. Tian, C.-W. Wong, and M. Wu, "Learning your heart actions from pulse: ECG waveform reconstruction from PPG," *IEEE Internet Things J.*, vol. 8, no. 23, pp. 16734–16748, Dec. 2021, doi: 10.1109/JIOT.2021.3097946.

[13] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, Jun. 2000.

[14] M. Kachuee, M. M. Kiani, H. Mohammadzade, and M. Shabany, "Cuffless high-accuracy calibration-free blood pressure estimation using pulse transit time," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2015, pp. 1006–1009.

[15] M. Saeed, C. Lieu, G. Raber, and R. G. Mark, "MIMIC II: A massive temporal ICU patient database to support research in intelligent patient monitoring," in *Proc. Comput. Cardiol.*, 2002, pp. 641–644.

[16] T. Athaya and S. Choi, "An efficient fingertip photoplethysmographic signal artifact detection method: A machine learning approach," *J. Sensors*, vol. 2021, pp. 1–18, Oct. 2021.

[17] M. Elgendi, "Optimal signal quality index for photoplethysmogram signals," *Bioengineering*, vol. 3, no. 4, p. 21, 2016.

[18] N. Gambarotta, F. Aletti, G. Baselli, and M. Ferrario, "A review of methods for the signal quality assessment to improve reliability of heart rate and blood pressures derived parameters," *Med. Biol. Eng. Comput.*, vol. 54, no. 7, pp. 1025–1035, Jul. 2016.

[19] S. Huthart, M. Elgendi, D. Zheng, G. Stansby, and J. Allen, "Advancing PPG signal quality and know-how through knowledge translation-from experts to student and researcher," *Frontiers Digit. Health*, vol. 2, pp. 1–7, Dec. 2020.

[20] E. K. Naeini, I. Azimi, A. M. Rahmani, P. Liljeberg, and N. Dutt, "A real-time PPG quality assessment approach for healthcare Internet-of-Things," *Proc. Comput. Sci.*, vol. 151, pp. 551–558, Jan. 2019.

[21] D. Roh and H. Shin, "Recurrence plot and machine learning for signal quality assessment of photoplethysmogram in mobile environment," *Sensors*, vol. 21, no. 6, p. 2188, Mar. 2021.

[22] J. Sun, A. Reisner, and R. Mark, "A signal abnormality index for arterial blood pressure waveforms," in *Proc. Comput. Cardiol.*, 2006, pp. 13–16.

[23] M. Elgendi, R. Fletcher, Y. Liang, N. Howard, N. H. Lovell, D. Abbott, K. Lim, and R. Ward, "The use of photoplethysmography for assessing hypertension," *NPJ Digit. Med.*, vol. 2, no. 1, pp. 1–11, Dec. 2019.

[24] M. Kachuee, M. M. Kiani, H. Mohammadzade, and M. Shabany, "Cuffless blood pressure estimation algorithms for continuous health-care monitoring," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 4, pp. 859–869, Apr. 2017.

[25] L. Peter, N. Noury, and M. Cerny, "A review of methods for non-invasive and continuous blood pressure monitoring: Pulse transit time method is promising?" *IRBM*, vol. 35, no. 5, pp. 271–282, Oct. 2014.

[26] F. Shirbani, C. Blackmore, C. Kazzi, I. Tan, M. Butlin, and A. P. Avolio, "Sensitivity of video-based pulse arrival time to dynamic blood pressure changes," in *Proc. 40th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2018, pp. 3639–3641.

[27] S. G. Khalid, J. Zhang, F. Chen, and D. Zheng, "Blood pressure estimation using photoplethysmography only: Comparison between different machine learning approaches," *J. Healthcare Eng.*, vol. 2018, pp. 1–13, Oct. 2018.

[28] Y. Liang, Z. Chen, G. Liu, and M. Elgendi, "A new, short-recorded photoplethysmogram dataset for blood pressure monitoring in China," *Sci. Data*, vol. 5, no. 1, pp. 1–7, Dec. 2018.

[29] S. Asgari, M. Bergsneider, and X. Hu, "A robust approach toward recognizing valid arterial-blood-pressure pulses," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 1, pp. 166–172, Jan. 2010.

[30] Q. Li and G. D. Clifford, "Dynamic time warping and machine learning for signal quality assessment of pulsatile signals," *Physiol. Meas.*, vol. 33, no. 9, p. 1491, Sep. 2012.

[31] T. Athaya and S. Choi, "Evaluation of different machine learning models for photoplethysmogram signal artifact detection," in *Proc. Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, Oct. 2020, pp. 1206–1208.

[32] O. Zhang, C. Ding, T. Pereira, R. Xiao, K. Gadhoumi, K. Meisel, R. J. Lee, Y. Chen, and X. Hu, "Explainability metrics of deep convolutional networks for photoplethysmography quality assessment," *IEEE Access*, vol. 9, pp. 29736–29745, 2021.

[33] H. Hsiu, C.-L. Hsu, and T.-L. Wu, "A preliminary study on the correlation of frequency components between finger PPG and radial arterial BP waveforms," in *Proc. Int. Conf. Biomed. Pharmaceutical Eng.*, Dec. 2009, pp. 1–4.
[34] E. J. Argüello-Prada, "The mountaineer's method for peak detection in photoplethysmographic signals," *Revista Facultad de Ingeniería Universidad de Antioquia*, vol. 90, pp. 42–50, Jan. 2019.
[35] C. Fischer, M. Glos, T. Penzel, and I. Fietze, "Extended algorithm for real-time pulse waveform segmentation and artifact detection in photoplethysmograms," *Somnologie*, vol. 21, no. 2, pp. 110–120, Jun. 2017.
[36] M. Salah, L. Hassan, S. Abdel-khier, A. M. Hassan, and O. A. Omer, "Robust facial-based inter-beat interval estimation through spectral signature tracking and periodic filtering," in *Intelligent Sustainable Systems*. Singapore: Springer, 2022, pp. 161–171.
[37] N. Yoder, "Peakfinder (x0, sel, thresh, extrema, includeEndpoints, interpolate)," MATLAB Central File Exchange, Tech. Rep., 2021.
[38] G. Tusman, C. M. Acosta, S. Pulletz, S. H. Böhm, A. Scandurra, J. M. Arca, M. Madorno, and F. S. Sipmann, "Photoplethysmographic characterization of vascular tone mediated changes in arterial pressure: An observational study," *J. Clin. Monitor. Comput.*, vol. 33, no. 5, pp. 815–824, Oct. 2019.
[39] T. Abhay, "Estimating correlation between arterial blood pressure and photoplethysmograph," in *Proc. 16th Int. Conf. Biomed. Eng.*, 2017, pp. 47–52.
[40] L. N. Harfiya, C.-C. Chang, and Y.-H. Li, "Continuous blood pressure estimation using exclusively photopletysmography by LSTM-based signal-to-signal translation," *Sensors*, vol. 21, no. 9, p. 2952, Apr. 2021.
[41] I. T. Jolliffe and J. Cadima, "Principal component analysis: A review and recent developments," *Phil. Trans. Roy. Soc. A*, vol. 374, Apr. 2016, Art. no. 20150202.
[42] K. R. Gabriel, "The biplot graphic display of matrices with application to principal component analysis," *Biometrika*, vol. 58, no. 3, pp. 453–467, 1971.
[43] H. Sak, A. Senior, and F. Beaufays, "Long short-term memory based recurrent neural network architectures for large vocabulary speech recognition," 2014, *arXiv:1402.1128*.

MOSTAFA SALAH received the B.Sc. and M.Sc. degrees in electrical engineering and the Ph.D. degree in communication engineering from Assiut University, Assiut, Egypt, in 2003, 2010, and 2018, respectively. Since 2018, he has been an Assistant Professor of electronics and communications engineering with the Faculty of Engineering, Sohag University. His research interests include mobile wireless communications, metasurface applications, embedding super-resolution and compressive sensing in single/multicarrier modulation, indexed modulation, and aggregation of millimeter wave with cellular mobile systems.

OSAMA A. OMER (Member, IEEE) received the B.Sc. and M.Sc. degrees from South Valley University, in 2000 and 2004, respectively, and the Ph.D. degree from the Tokyo University of Agriculture and Technology, in 2009. He has spent six month as a Postdoctoral Researcher with the Medical Engineering Department, Luebeck University, Germany. He has also spent three months as a Postdoctoral Researcher at Kyushu University, Japan. The last but not least, he has spent six months as a Research and Development Scientist Engineer at the NOKIA Research and Development Center, Tokyo/Japan, in 2008. He is currently a Full Professor at the Faculty of Engineering, Aswan University. His research interests include wireless communications, deep learning, and image/signal processing.

LOAY HASSAN received the B.Sc. and M.Sc. degrees from Aswan University, in 2010 and 2015, respectively, where he is currently pursuing the Ph.D. degree. He has worked as an Assistant Researcher at the Center for Artificial Intelligence and Robotics (CAIRO), Aswan University, for four years. He has worked as a part-time Teacher Assistant at the Department of Computer Science, Arab Academy for Science, Technology and Maritime Transport, South Valley Branch. He is currently a Researcher in computer engineering at University Rovira i Virgili (URV).

MOHAMED RAGAB received the B.Sc. (Hons.) and M.Sc. degrees from the Department of Electrical Engineering, Aswan University, in 2014 and 2017, respectively. He is currently pursuing the Ph.D. degree with the School of Computer Science and Engineering, Nanyang Technological University (NTU), Singapore. Concurrently, he is with the Machine Intellection (MI) Department, Institute for Infocomm Research (I2R), A*STAR. His research interests include deep learning, transfer learning, and intelligent fault diagnosis and prognosis.

AMMAR MOSTAFA HASSAN received the B.Sc. and M.Sc. degrees from Assiut University, Asyut, Egypt, in 1995 and 2005, respectively, and the Ph.D. degree from Minia University, Minya, Egypt, in 2012. He has spent two years at Otto-von-Guericke University, Magdeburg, Germany. He is currently the Head of the Computer Science Department, Arab Academy for Science, Technology, and Maritime Transport, South Valley Branch. His research interests include image processing, image hashing, cryptography, automatic/digital control, optimization, computer security, and image authentication.

AHMED ABDELREHEEM received the B.Sc. and M.Sc. degrees and the Ph.D. degree in communication engineering from Aswan University, in 2004, 2015, and 2019, respectively. He is currently an Assistant Professor with the Information Technology Department, The National Egyptian Electronic Learning University. His research interests include medical imaging, deep learning, 6G networks, and terahertz communications.