



# Video-based beat-by-beat blood pressure monitoring via transfer deep-learning

Osama A. Omer<sup>1</sup> · Mostafa Salah<sup>2</sup> · Loay Hassan<sup>1</sup> · Ahmed Abdelreheem<sup>3</sup> · Ammar M. Hassan<sup>4</sup>

Accepted: 20 February 2024 / Published online: 1 April 2024  
© The Author(s) 2024

## Abstract

Currently, learning physiological vital signs such as blood pressure (BP), hemoglobin levels, and oxygen saturation, from Photoplethysmography (PPG) signal, is receiving more attention. Despite successive progress that has been made so far, continuously revealing new aspects characterizes that field as a rich research topic. It includes a diverse number of critical points represented in signal denoising, data cleaning, employed features, feature format, feature selection, feature domain, model structure, problem formulation (regression or classification), and model combinations. It is worth noting that extensive research efforts are devoted to utilizing different variants of machine learning and deep learning models while transfer learning is not fully explored yet. So, in this paper, we are introducing a per-beat rPPG-to-BP mapping scheme based on transfer learning. An interesting representation of a 1-D PPG signal as a 2-D image is proposed for enabling powerful off-the-shelf image-based models through transfer learning. It resolves limitations about training data size due to strict data cleaning. Also, it enhances model generalization by exploiting underlying excellent feature extraction. Moreover, non-uniform data distribution (data skewness) is partially resolved by introducing logarithmic transformation. Furthermore, double cleaning is applied for training contact PPG data and testing rPPG beats as well. The quality of the segmented beats is tested by checking some of the related quality metrics. Hence, the prediction reliability is enhanced by excluding deformed beats. Varying rPPG quality is relaxed by selecting beats during intervals of the highest signal strength. Based on the experimental results, the proposed system outperforms the state-of-the-art systems in the sense of mean absolute error (MAE) and standard deviation (STD). STD for the test data is decreased to 5.4782 and 3.8539 for SBP and DBP, respectively. Also, MAE decreased to 2.3453 and 1.6854 for SBP and DBP, respectively. Moreover, the results for BP estimation from real video reveal that the STD reaches 8.027882 and 6.013052 for SBP and DBP, respectively. Also, MAE for the estimated BP from real videos reaches 7.052803 and 5.616028 for SBP and DBP, respectively.

**Keywords** Blood pressure · Remote photoplethysmography (rPPG) · Unobtrusive monitoring · Skin segmentation · Beat segmentation · ImageNet · 2D Beats

## Abbreviations

Abbreviations	Description		
ABP	Arterial Blood Pressure	cPPG	contact Photoplethysmography
AI	Artificial Intelligence	CQI	Beat Correlation Quality Index 5
BbB	Beat-by-Beat	DBP	Diastolic Blood Pressure
BI	Beat Interval	DL	Deep Learning
BP	Blood Pressure	LIME	Locally Interpretable Model Agnostic Explanations
CNN	Convolutional neural network	MAB	Mean Arterial Blood Pressure
		PPG	Photoplethysmography
		PPT	Pulse Transit Time
		PWV	Pulse Wave Velocity
		rPPG	remote Photoplethysmography
		ROI	Region-Of-Interest
		SBP	Systolic Blood Pressure
		SHAP	SHapley Additive exPlanation

✉ Mostafa Salah  
mostafa.salah@eng.sohag.edu.eg

Extended author information available on the last page of the article

SQI	Beat Skewness Quality Index
XAI	Explainable Artificial Intelligence

## 1 Introduction

Abnormal blood pressure (BP) is one of the most critical biomarkers for cardiovascular diseases that are ranked as the most common causes of death worldwide as reported by the world health organization (WHO) [1]. BP measurement can be classified into invasive and non-invasive methods. The invasive method provides continuous arterial blood pressure (ABP) monitoring needed for patients in the intensive care unit (ICU) or under high-risk surgery. It is performed via arterial cannulation by a trained operator. The invasive method represents the gold standard of BP monitoring, however, it is difficult to measure in routine clinical practice along with some pitfalls as addressed in [2]. On the other hand, the conventional non-invasive methods can be classified according to 1) the monitoring rate as continuous or intermittent, and, 2) triggering as automated or manual [3]. Traditional non-invasive methods include auscultatory, oscillometry, tonometry, and volume clamping techniques [4]. These techniques vary in reliability levels and usability. Some of these techniques are not always suitable for babies and elder people because the measurement includes inflation of the cuff to a pressure above the systolic pressure. A proper BP monitoring method should be selected according to the patient's case and the required monitoring rate. For example, a simple cuff-based method can be used by the medical staff as a routine test in clinics. However, a higher monitoring rate (every five minutes at least) is required for patients under anesthesia according to the recommendations of the American Society of Anesthesiologists (ASA) [5] for checking the functionality of the patient's circulatory. Motivated by the current trend of increased health awareness, the need arises for ubiquitous sensing and monitoring of critical biomarkers continuously [6, 7]. Simple and continuous non-clinical BP monitoring becomes highly desirable nowadays for limiting the risk of cardiovascular diseases and hypertension. It is not limited to just monitoring tasks, it is extended to diagnosing and predicting the problems in advance [8]. Consequently, considerable research efforts have been recently conducted on finding comfortable and accurate means for BP self-monitoring constantly in a non-clinical environment [9, 10]. In this context, photoplethysmography (PPG) signal [11, 12] exhibits an essential role in the non-invasive and continuous monitoring of many vital signs [13] such as heart rate variability [14], respiration rate [15], blood pressure [16, 17], electrocardiogram (ECG) reconstruction [18–20], hemoglobin level [21], and oxygen saturation level (SpO<sub>2</sub>) [22]. Motivated by PPG extraction simplicity, many wearable devices [23–26] are introduced for predicting these

physiological vital signs from the PPG signal that represents changes in the volume of the blood inside the arteries due to heart pulsation. Fortunately, the corresponding changes in blood volume can be inferred by regarding tiny changes in skin color through the RGB camera [27–33]. That way refers to the imaging PPG (iPPG) or remote-PPG (rPPG) technique employed for providing contactless PPG extraction. The wide availability of smartphones empowers rPPG-based vital sign detection [34–41] without additional hardware. Based on the extracted PPG/rPPG, many vital signs can be extracted from the measured blood volume changes in skin vessels. Thanks to publically available biomedical datasets [42–45], machine/deep learning techniques [46] can be applied for inferring the underlying relationship between the shape of the PPG signal and other vital signs.

By turning on our main interest in BP estimation from PPG/rPPG signal, extracting a clean facial rPPG signal (with an accepted signal-to-noise ratio) is an extremely challenging operation. That point has many problematic details [47–50] such as light variation, motion artifacts, camera grade, video coding, surrounding interference, and heart rate modulation by respiration rate. So, many researchers are working on rPPG-based BP estimation from different perspectives. Despite hundreds of related papers [46, 51] and emerging benchmark reporting about rPPG extraction and deep/machine learning-based BP prediction [52–54], uncovering new features is a non-stopping operation [55–57]. Also, it has many issues and challenges [58]. Many researchers strike the problem from different perspectives including signal de-noising, defining the region of interest (ROI), data cleaning, feature selection, feature combinations, feature domain, feature format, model selection, and model personalization. So, the research topic has a big picture with increasing sides. In this paper, we are interested in resolving some problems related to the nature of the training data (data shortage after strict cleaning and non-uniform labeling distribution) and the reduced quality of the extracted rPPG signal. Section 2 gives a fast literature review for both rPPG extraction and learning models. We have emphasized our main interest which is related to transfer learning. Section 3, the addressed challenges and the introduced contribution is provided in detail. Section 4 introduces a detailed description of our proposed system including video/signal processing, training, and testing stages. Experimental results are demonstrated in Section 5 based on both contact PPG and real videos. Finally, Section 6 presents the conclusion of this work.

## 2 Literature review

The literature review can be divided essentially into two main parts including rPPG extraction and PPG-BP training models.

## 2.1 Camera-based rPPG extraction overview

As reported, the PPG signal can be estimated remotely from a facial video stream [28, 30, 31, 54]. The tiny light variation of the illuminated facial skin is captured by the RGB or infrared camera [30, 59] where the reflected light is modulated by the changing amount of red blood cells in skin vessels due to cardiovascular pulsation. Both PPG and rPPG signals carry similar information. However, extracting rPPG is a more challenging task than contact PPG extraction. There are many sources of noise and interference [49, 50] such as illumination instability [47], low light environment [60], subject motion, and small head vibration [40, 48]. Also, the employed video coding (compression) has an essential impact on rPPG quality [61]. Aggressive compression may destroy the signal signature completely. While uncompressed video preserves most of the signal features, it suffers from the need for huge memory. Furthermore, there are many variations in the camera grade and capture settings [62, 63]. For example, expensive cameras and special capturing settings are reported [64] for providing higher signal quality. On the other hand, rPPG extraction is more convenient than contact PPG because it is contactless and easy to use without special devices where the cameras are essentially available on smartphones. Therefore, many researchers emphasize the consumer-level camera grade [65, 66] or smartphone camera [34, 35, 39, 40] for allowing cost-effective vital sign solutions. Furthermore, the camera-based approach helps in monitoring the BP many times and may be continuously used without side effects or pain. The rPPG extracting involves more sophisticated image and signal processing stages [54]. The successive processing stages can be explained as follow:

- **Face detection and tracking:** For ensuring proper signal extraction, the face should be detected and tracked continuously through the video frames.
- **Regions of interest (ROI) detection:** Some ROI exhibits higher SNR levels. Hence, it can be localized for providing enhanced signal strength.
- **Skin segmentation:** Moreover, skin segmentation can be applied to catch skin pixels only without any background pixels [67, 68].
- **1-D signal conversion:** The video sequence is converted into a 1-D signal by averaging the intensity levels of the whole frame pixels or the localized pixels in the intended ROI [69]. So, we have three independent vectors arising from RGB channels. Among the RGB channels, the green channel is preferred because it conveys the highest signal power [28, 50, 54].
- **Filtering and de-noising:** To that end, the green channel signal is just raw data that needs further filtering and

denoising. Many filtering approaches are introduced [28, 65], however, rPPG denoising remains an essential open problem [70].

- **Beat segmentation:** The rPPG signal is divided into pulsating beats [71–74].
- **BP prediction Model:** The filtered rPPG signal is fed to the trained model in beat or signal level. The input should be normalized in the amplitude. The length should be resized into a fixed number of samples.

## 2.2 Training/non-training models

By turning on our main interest in BP estimation from PPG/rPPG signal, an interesting classification of the applied training approaches can be found in [53] where the related work is divided into parameterized methods, and data-driven models that include both machine learning features, and deep learning (end-to-end prediction).

**1) Parameterized methods:** The parameterized methods involve pulse arrival time (PAT) and pulse transit time (PTT). The PTT represents the pulse traveling time between two arterial sites [75, 76]. It needs monitoring PPG signals from two different sensors located at a known distance. So it is not preferred compared to single sensor methods. On the other hand, PAT is related to a certain time shift between PPG and ECG signals [77]. However, it needs to extract the ECG signal besides the PPG signal. Using two kinds of signals, ECG and PPG signals, in this method represents a weak point. Moreover, the usually used Medical Information Mart for Intensive Care (MIMIC) dataset does not guarantee synchronization between acquired PPG and ECG signals [78].

**2) Data-driven methods:** Although the employed dataset is common and most of the employed models are already known, there are an increasing number of learning-based papers concerning blood pressure estimation from PPG signals because there is a diverse number of critical points represented in data cleaning, employed features, feature format, feature selection, feature domain, model structure, problem formulation (regression or classification) and model combinations. Hence, it becomes necessary to provide general guidelines for a fair assessment of any claimed results [56]. There are many excellent comprehensive surveys [51, 79] regarding both machine learning and deep learning-based BP estimation. So, in this section, we provide only some examples of these methods. It is worth noting that almost all the related work exploits machine learning and deep learning approaches while very limited publications utilize a transfer learning approach that represents our main interest.

**A. Machine learning and deep learning:** The main advantage of a data-driven approach resides in relying on a single PPG signal without the need for additional signals.

For example, only morphological PPG shape is employed for extracting some training features such as relative amplitudes/intervals between certain points on the beat [80–82]. Furthermore, spectral features can be combined with temporal features for predicting BP [83–85]. The impact of data cleaning is highlighted in [86, 87]. The problem may be formulated as a regression or classification problem [88]. BP classification can be simply divided into three classes [16, 85] namely, normotensive (NT), prehypertensive (PHT), and hypertensive (HT). Regression may be preceded by classification of PPG beats for enhancing BP estimation accuracy [89]. Deep neural networks can be used for learning latent features [90–92]. The network depth was studied also [93]. Complete ABP signal may be mapped from the corresponding PPG signal through signal level transformation [94, 95]. Then, ABP has been eventually used to provide systolic blood pressure (SBP) and diastolic blood pressure (DBP) by finding the maximum and minimum values of the ABP signal, respectively. However, any signal quality degradation for a small interval may degrade the prediction accuracy. Hence, it is recommended to work on a beat-bases where the distorted beats can be neglected. Other than PPG-ABP signal translation, simple systolic and diastolic BP prediction is sufficient in most cases. In [96], federated Learning is employed for enabling distributed learning of ABP time series from PPG time series. Although it has a very low mean absolute error, the standard deviation is relatively high.

**B. Transfer learning:** In transfer learning, the first layers provide general (not limited to specific datasets) feature learning where the neural network was trained on a huge dataset such as ImageNet dataset containing natural images. The learned feature capability can be extended to our target dataset by fine-tuning the last layer [97]. Hence, it improves learning performance on relatively small datasets. However, the main differences reside in transfer purposes and finding an efficient representative way for mapping a 1-D PPG signal into a 2-D image. Despite the huge amount of research efforts devoted to deep learning-based BP estimation, there are a very limited number of publications at that point. For instance, the 1-D PPG signal is converted into a 2-D image through the transform domain (spectrogram and scalogram) for providing fine-tuning of ResNet-18 [98]. In [99], the visibility graph is used for creating corresponding 2-D images. Then, the pre-trained deep convolutional network (CNN) is exploited for feature extraction in the first layers. Other than 2-D transfer learning, the 1-D pre-trained for one patient may be transferred to another patient by utilizing learned parameters in the first layers. Hence, avoiding learning from scratch and improving model generalization [100] between similarly clustered signals. Also, generally trained models can be personalized by exploiting performed training on the whole PPG-BP dataset to be tuned on the intended subject [101]. This reduces the number of required samples for new

subject personalization and improves learning performance [102]. On the other hand, performed classification and regression may be refined by giving a higher impact on some of the strongly related features [103].

### 3 Challenges and contributions

Based on the foregoing, there is a growing interest in camera-based BP estimation from research and industrial perspectives. However, there are still many challenges. In this paper, we shed light on some of these challenges and propose some recommended solutions as follows:

#### 3.1 Challenges and limitations

**1) Reduced amount of data after cleaning:** Even though contact PPG is regarded as more reliable than remotely extracted PPG, a huge amount of contact PPG is susceptible to many sources of interference and distortion. For ensuring well-trained deep learning models, the learning dataset should be strictly cleaned by dropping outliers and noisy/deformed signals. Different quality assessment strategies are applied along with proper thresholds for excluding unaccepted signals/beats [104–110]. However, strict cleaning reduces the training data severely. For example, the work presented in [83] applies strict cleaning on MIMIC III data [43] which leads to reducing valid data from 30000 patient records down to 510 records only. Consequently, the amount of evidently cleaned data is reduced significantly. It demands 10 min long record at least with high SNR quality.

**2) The reduced rPPG extraction quality:** The extracted rPPG signal (corresponding to the tiny variations of blood volume in the facial skin tissue) has a very low signal-to-noise ratio (SNR) and changes over time. There are many sources of noise and interference (video compression, motion, and light variation). So, the extraction quality varies from beat to beat. BP estimation based on the distorted signal increases prediction errors. It is predicted that camera-based remote PPG is more challenging and highly susceptible to interference and distortion.

**3) Skewed data histogram:** The BP labeling range is not covered equally by the available CPPG data. The high-pressure range has a lower representation than other BP ranges. The BP distribution is skewed. It introduces some biasing or a preferred prediction range.

**4) Learning schemes:** Extensive research efforts are devoted to utilizing either deep learning or machine learning [46, 51] for training models. However, despite the availability of well-known off-the-shelf image-based models, there is very limited work on exploiting transfer learning in BP prediction from rPPG/CPG signals. Most of that work was limited mainly to personalizing [100] the generalized trained

models without exploiting its impact on providing more generalized models.

### 3.2 Contributions

The main contributions of this work can be summarized in the following:

- Smart per-beat cPPG cleaning: Long PPG records may exhibit low SNR due to the deformation of some individual beats or during short intervals. Hence, the whole record may be regarded as noisy and dropped from training operations. Other than long-stream PPG cleaning/training [83, 94, 95] for PPG signal-to-ABP signal transformation, we follow smart double cleaning [87] where the coarse cleaning is applied on the signal level first, then, fine cleaning is applied on the beat level. Hence, deformed beats can be dropped individually without losing the whole record.
- Per-beat rPPG selection: Many denoising schemes are introduced [28, 65, 111–113], however, providing long streams of high-quality rPPG signals still remains a challenging task. While it is difficult to extract long stream rPPG signal with high quality, it is more proper to select non-deformed beats during high SNR intervals. Hence, instead of complete signal prediction as in [94], we provide a per-beat BP estimation approach where the extracted rPPG is segmented into beats and the deformed beats are excluded from the BP prediction stage based on some quality measures of the rPPG/PPG signals. The intervals of low rPPG quality can be ignored while the estimation decision is built on more trusted beats during the best instances. Also, aside from PPG/ABP signal prediction, per-beat BP monitoring introduces a trusted high monitoring rate where individually catching non-
- distorted beats is more probable than catching complete regular long segments (multiple beats in a successive stream). There are no single selection criteria for quality estimation. So, we propose to utilize some successive selection rules for selecting optimal beats for BP estimation.
- Logarithmic labeling: For partially resolving data skewness, the real BP labeling (systolic and diastolic) is transformed into a new range of labeling through logarithmic labeling. It will be demonstrated that the results of the logarithmic labeling outperform the results of the original labeling.
- Transfer learning Review: As it is our main interest, we introduce a detailed review of the applied transfer learning for BP estimation from PPG signal.
- Beat-to-image enabled Transfer Learning: For alleviating the drop in the amount of data after data cleaning and enhancing model generalization with limited data, we rely on transfer learning. The image-based pre-trained deep learning networks are exploited as a starting point for the PPG-BP training task. The 1-D beat is converted into a 2-D image where the temporal amplitude adapts to the intensity of the pixels. PPG morphology and beat interval are preserved in the created image. furthermore, five deep learning networks are tested including AlexNet, Resnet101, VGG16, MobileNet, and DenseNet. We note that the Resnet101 network outperforms other tested networks.
- The 2-D training is compared to the 1-D training.

## 4 Proposed blood pressure monitoring

As shown in Fig. 1, the overall camera-based BP assessment system is explained. It depends on the detection of remote

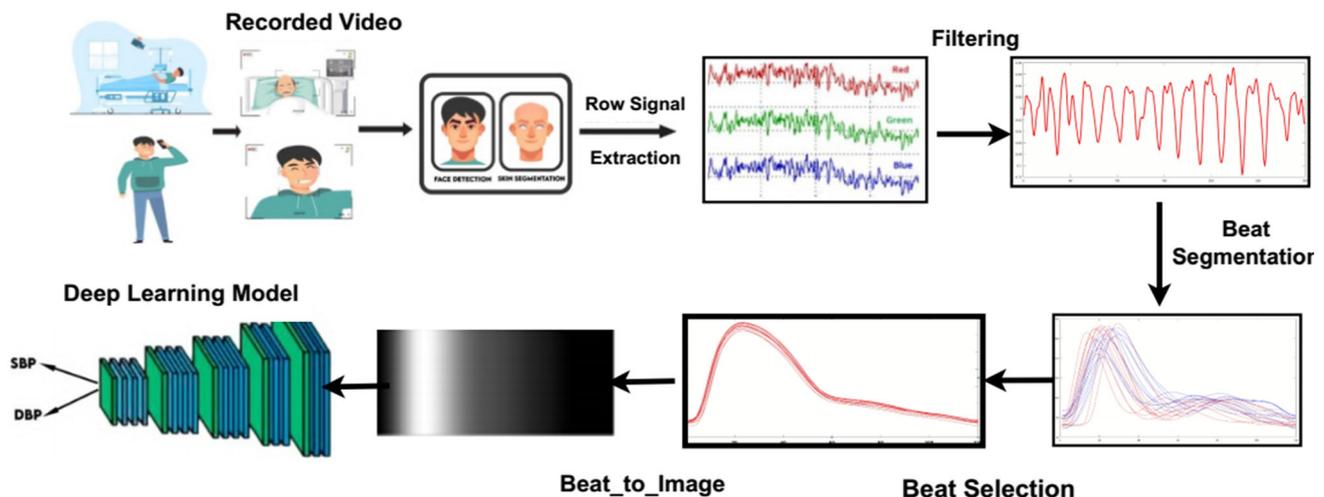


Fig. 1 Proposed camera-based BP assessment system

rPPG signal which is eventually used to estimate BP. The proposed system consists of four steps. In the first step, the raw data are extracted from the video stream after a facial skin segmentation. Then, the raw data undergoes pre-processing operations including multi-band filtering for denoising the rPPG signal. Then, the rPPG signal is segmented into beats where the beat selection scheme is applied to reject the distorted beats. Selected beats are therefore converted into 2D images. The beat intervals are implicitly included together with the rPPG beats in the 2D images. In the last step, the selected beats are used for BP prediction utilizing deep learning networks. Figure 2 shows the block diagram of the suggested system. Details of the proposed system will be described in the following subsections. Specifically, the proposed system consists of two phases; namely, the learning phase and the testing/prediction phase.

## 4.1 Learning phase

### 4.1.1 Dataset

Although we are introducing a deep-learning model for mapping the rPPG signal (video-based) to BP, it is common to use the contact PPG (CPPG) dataset for the learning phase for the following reasons: 1) Compared to CPPG, the available video-based rPPG dataset is too small [114–116]; 2) rPPG has many different capturing settings (illumination and camera grade); 3) The rPPG is much noisier than the corresponding

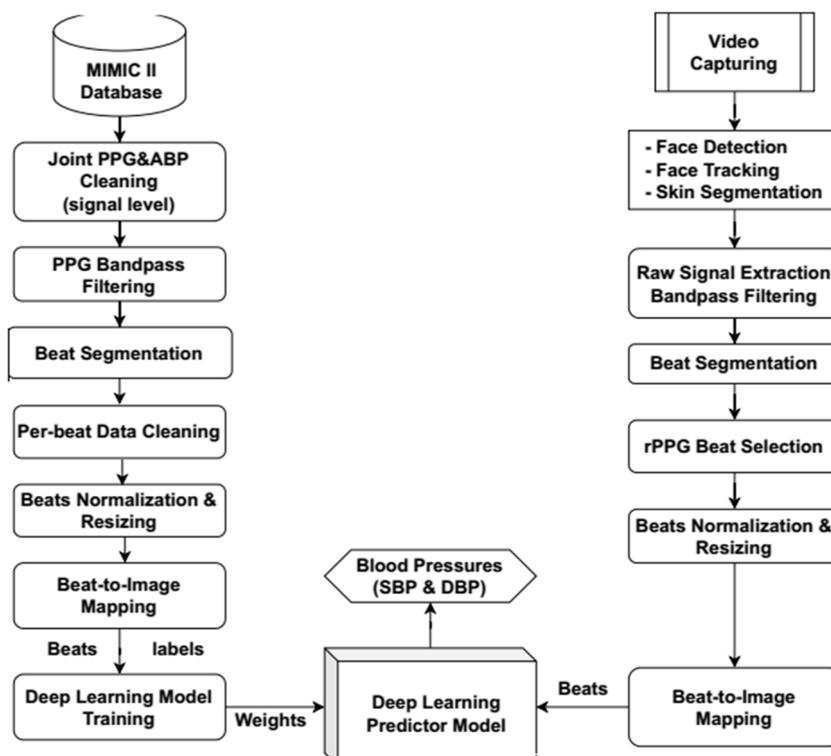
contact PPG, 4) Both rPPG and contact PPG share the same characteristics; 5) There is a huge amount of contact dataset along ground-truth values [42–45, 86, 117–119]. Generally, available data sets cannot be utilized directly [120]. Extensive pre-processing/cleaning is needed before any practical usage. However, the MIMIC dataset represents the largest available dataset. There are many versions of MIMIC such as MIMIC-I, MIMIC-II, and MIMIC-III [42, 43, 45]. Newer versions include more subject measurements and some updates of data structures [121]. In this paper, we prefer to operate on MIMIC-II. While the MIMIC dataset represents the universal common source of the dataset for almost all the related work, the essential difference resides in the applied cleaning strategy [86, 117]. However, we follow our cleaned version [87].

In the learning phase, the dataset passes through the following processing steps.

### 4.1.2 PPG cleaning

The data contains many distorted signals that should be excluded before any training. The original MIMIC II dataset contains independent records with different lengths. The corresponding PPG and ABP signals are segmented into lower intervals for proper processing. Successive cleaning metrics can be applied for excluding improper signals [87]. The cleaning is performed at the signal level and the beat level as well. At the signal level, the spectral construction of both

**Fig. 2** The block diagram of the proposed system (training and testing stages)



signals is determined. Clear signals exhibit high similarity in the frequency domain because they arise from the same pulsating source (heart). Based on the signal periodicity, both signals should agree on the fundamental frequency (heart rate) and the related harmonics. Heart rate mismatching represents a strong sign of noisy signals. Also, detecting heart rate out of our physiological limits is another sign of noisy and distorted signals. Based on the known spectral shape and periodicity of PPG/ABP signals, signal power spectral density is concentrated around the fundamental frequency and its harmonics with a very narrow bandwidth (about 0.2 Hz). So, another metric can be tested by checking the spectral concentration through the signal-to-noise ratio (SNR) of both PPG and ABP signals which can be determined as the ratio of the power of the interesting signal (around fundamental and harmonics) to the power of the out-band representing the noisy component:

$$Q_{SNR} = \frac{\int_{f_0-\Delta f}^{f_0+\Delta f} \hat{P}_f df + \int_{2f_0-\Delta f}^{2f_0+\Delta f} \hat{P}_f df + \int_{3f_0-\Delta f}^{3f_0+\Delta f} \hat{P}_f df}{\int_{\Omega} \hat{P}_f df - [\int_{f_0-\Delta f}^{f_0+\Delta f} \hat{P}_f df + \int_{2f_0-\Delta f}^{2f_0+\Delta f} \hat{P}_f df + \int_{3f_0-\Delta f}^{3f_0+\Delta f} \hat{P}_f df]} \tag{1}$$

Where  $\hat{P}_f$  represents spectral power density measured over the cardiac band and  $\Omega = [0.8 - 5]$  Hz. The signals with SNR lower than a certain threshold are dropped out. Also, physiological limits are applied to the BP maximum value (Systolic BP) and minimum values (Diastolic BP). ABP signals having out-of-range BP are excluded too.

### 4.1.3 Signal filtering

The cleaned PPG signals are passed through two stages filtering process. In the first filter, the signals are passed through a bandpass filter in the cardiac frequencies [0.7 Hz - 5 Hz] to find the fundamental frequency. Usually, the fundamental frequency approximates the main signal periodicity and provides the average beat length. Based on the estimated fundamental frequency, the filtered signals are passed through a selective filter tuned at the fundamental frequency and its first harmonics for sharply identifying the beat intervals [38].

### 4.1.4 Beat segmentation

The filtered signal is segmented into beats to deal with each beat individually. The signals are segmented based on the detection of local minimum locations.

### 4.1.5 Per-beat data cleaning

The beat level represents the second cleaning level. Some cleaning metrics are applied successively on each beat. Based on the general distinct shape of PPG patterns, valid beats only

are maintained (selected) for training or testing stages. These metrics are beat intervals, skewness value, and correlation with the fundamental PCA component. Figure 3 provides an example of beat selection based on these metrics.

**Beat interval (BI):** Noisy signal incurs some errors in beat segmentation. So, we have improper beat length where the standard range of the heart rate is [40 bpm - 180 bpm] which corresponds to beat interval in the range [0.33 – 1.5] second. Therefore, we use only beats with beat intervals in the range  $0.33 \leq BI \leq 1.5$ . Also, we exclude beats with intervals too longer/shorter than the mean beat interval of successive beats.

**Beat skewness quality index (SQI):** Based on our observations, the normal (undisturbed) beats have a positively skewed shape. Also, it can be called the right-skewed beat. A tail is referred to as the tapering of the curve differently from the data points on the other side of the given beat is shifted to the right and with its tail on the left side, it is a negatively skewed beat. It is also called a left-skewed distribution. SQI can be calculated as

$$SQI = \frac{\sum_{i=1}^N (Y_i - \tilde{Y})^3 / N}{S^3} \tag{2}$$

Where  $\tilde{Y}$  is the mean,  $S$  is the standard deviation, and  $N$  is the number of beat's points

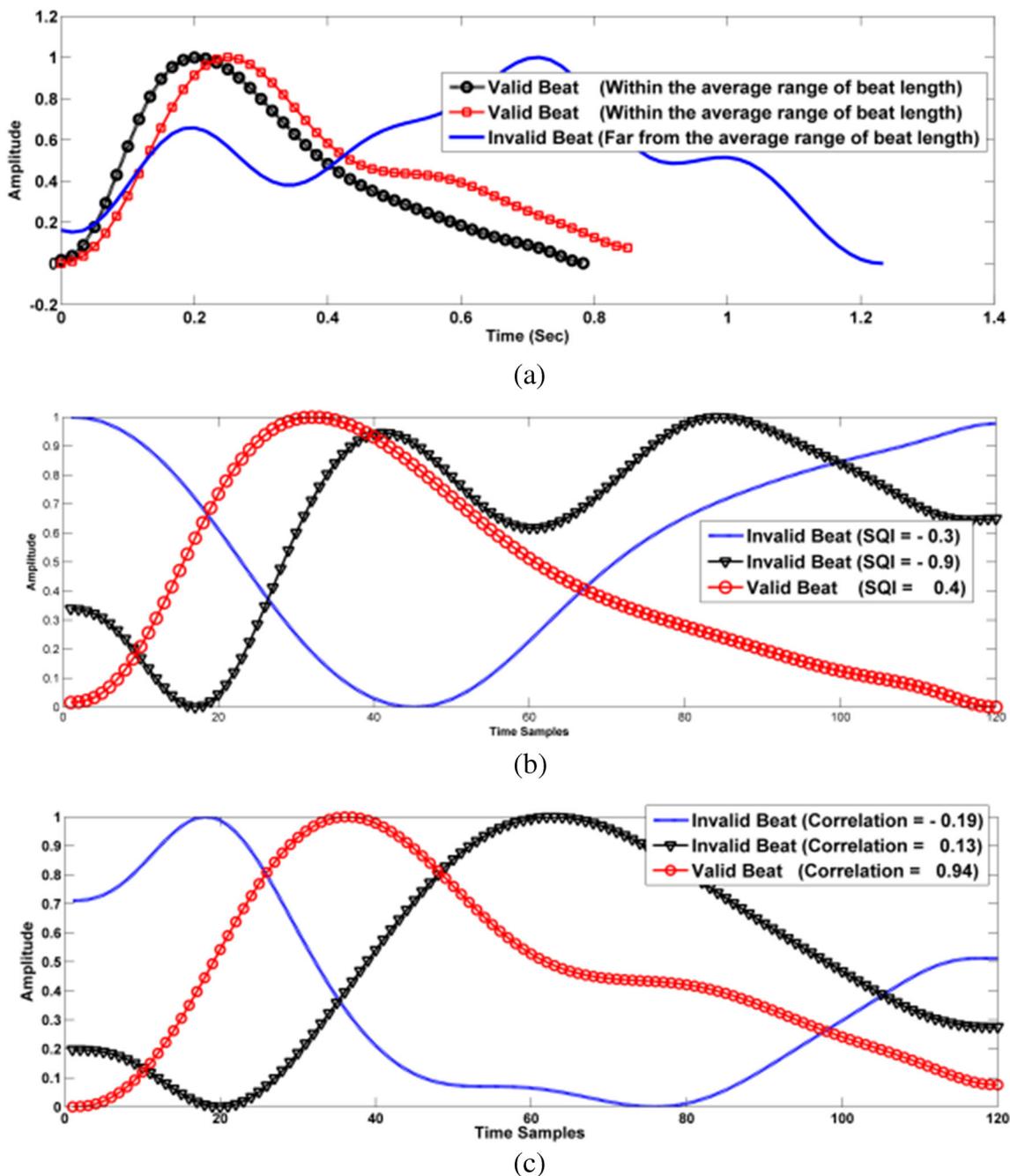
**Beat correlation quality index (CQI):** The segmented beats represent a new dataset that can be exploited for generating the essential building block of any beat through its first PCA component. Hence, the correlation with the fundamental PCA component can help for further beat cleaning by excluding beats that show poor correlation. However, the correlation should be not strictly specified. It is used to ensure the rejection of highly deviated beats only. At least, beats should have  $CQI > 0.3$ .

### 4.1.6 Beats normalization and resizing

The PPG signals are normalized in the amplitude to be in the range [0-1] by using the following equation.

$$S_n = \frac{S - \min(S)}{\max(S) - \min(S)}, \tag{3}$$

where  $S_n$  is the normalized signal and  $S$  is the un-normalized signal. The time interval is used implicitly as a feature besides the normalized PPG beats. The normalized beats are then normalized in time so as to be with fixed length (120) with PPG data representing  $BI \times \frac{120}{1.5}$  from the fixed length (120). The rest of the 120 samples are filled with zeros.



**Fig. 3** Example of Valid/Invalid beats for training/prediction phase (a) Time Interval metric. (b) Skewness quality index. (c) Correlation quality index

#### 4.1.7 Deep learning model training

For utilizing the pretrained image-based DL models such as ResNet101, VGG16, MobileNet, DenseNet, and AlexNet, segmented and selected beats are mapped into a 2-D image by varying the pixel intensity according to the beat amplitude as shown in Fig. 4. As the architectures of these networks are well known, an example of the network adaptation with our input is shown in Fig. 5 which is the ResNet101 architec-

ture. From this figure, we can see that the applied network is made up of the feature extractor part and the regressor. The feature extractor consists of one convolution receptive size of  $7 \times 7$  and a max pooling step followed by 4 blocks of similar behavior. Each of the blocks follows the same pattern. They perform  $1 \times 1$  convolution followed by  $3 \times 3$  convolution and another  $1 \times 1$  convolution with a fixed feature map dimension  $[64, 128, 256, 512]$  respectively. Furthermore, the width  $W$  and height  $H$  dimensions remain constant during the entire block.

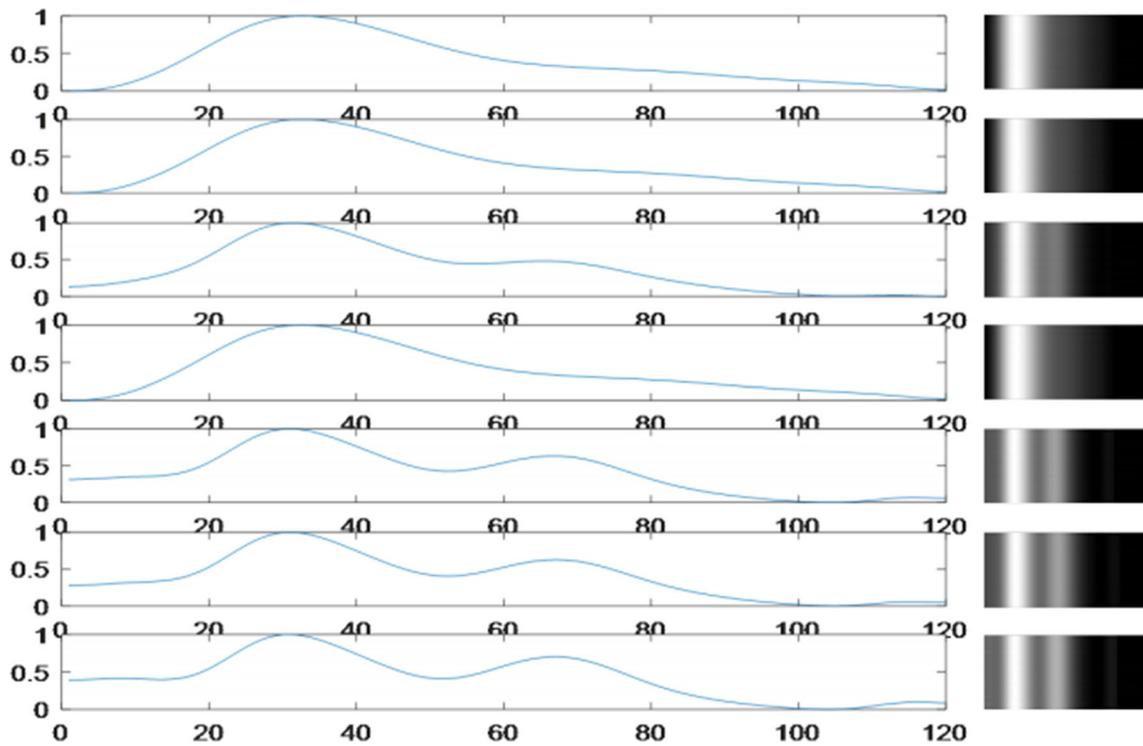


Fig. 4 1D Beat-to- 2D image mapping

The feature extractor is followed by two fully connected layers of the regressor with a flattening layer in between. The first one has 1000 neurons and the last fully connected layer serves as the output layer and has only one neuron to estimate the blood pressure value. In all the tested networks, the mean square error is utilized as a loss function in this network. For optimization, the ADAM optimizer is applied with a learning rate of  $1e-4$ . The networks are trained on 200 Epochs with a patch size of 20 images. Two scenarios are used with these networks. In the first scenario, we use normalized SBP and DBP values as labels for the networks. In the second scenario, we use normalized logarithmic SBP and logarithmic DBP values as labels for the networks to avoid distribution skewness.

### 4.2 Prediction phase

The input to this phase is the video stream. In this phase, the video stream passes through the following steps:

#### 4.2.1 Video analysis (face detection, tracking, and skin segmentation)

Deep learning networks are trained to detect and track the face and then a skin segmentation is applied. Face detection is considered a necessary first step for all facial analysis algorithms, including camera-based BP measuring. Recently, with the great success of deep learning in computer vision tasks, numerous deep learning-based detection algorithms

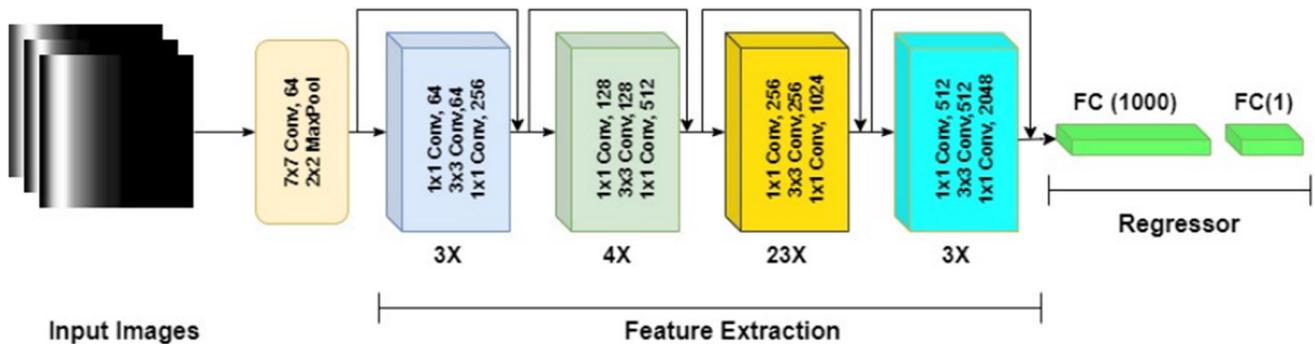


Fig. 5 Architecture of ResNet101

have been proposed such as MTCNN [122], Faster-RCNN [123], SSD [124], and YOLO [125]. Since measuring blood pressure by digital camera images requires speed and accuracy in extracting faces from the input images, in this work, we built our face detection algorithm based on YOLO v5 which has significant improvements compared to its predecessor versions. The face-tracking algorithm comes after the detection step to save time and computational complexity instead of running the face detection model on each frame from a digital video. Therefore, the high-speed and accurate KCF tracking algorithm [126] is used. On the other hand, the main target of the skin segmentation process is selecting the region of interest (ROI) within the detected face that contains pixels providing the raw RGB signal. Indeed, it is a critical process where the ROI must include as many skin pixels as likely for more accurate raw signal extraction. Hence a robust segmentation algorithm is needed. Several deep learning-based semantic segmentation models have been proposed that can be adapted for the skin segmentation task. In this work, SegNet Network [127] which is well-known for its high-performance deep learning-based segmentation model is employed. To improve the segmentation efficiency, unlike the original architecture proposed in [127] which consists of an encoder network that utilizes the VGG network [128] as a feature extractor and a decoder network followed by a pixel-wise classification layer, we replaced the VGG network on the encoder path with ResNet [129] as a backbone feature extractor for taking advantage of its powerful representational ability. We conducted the ablation study of the modified SegNet model to gain deep insight into the effects caused by replacing the VGG encoder network with the ResNet encoder network.

#### 4.2.2 Raw signal extraction

The average intensity of the skin pixels for each segmented frame is computed to produce a one-dimensional signal. This signal is a raw signal which needs further spectrum analysis and filtration to be ready for the rPPG estimation step.

#### 4.2.3 rPPG filtering and beat segmentation

The raw signal extracted from the previous step is filtered using two consecutive filters. The first filter is the bandpass filter with a pass band of [0.7 - 5] Hz. The second filter is the selective filter tuned at the fundamental frequency and its first harmonics. The fundamental frequency is estimated as the frequency that has the highest power [38].

#### 4.2.4 Best beats selection (rPPG cleaning)

Not all segmented beats are valid for BP prediction. Therefore, the same PPG dataset cleaning metrics in the learning

phase (described in Section 4.1) are also employed in the testing phase for valid rPPG beats selection (invalid beats rejection).

#### 4.2.5 Beats normalization and resizing

Selected rPPG beats are normalized to be in the range [0 - 1] and resized. Like the process done in the learning phase, the time interval is used implicitly as a feature besides the normalized beats. The normalized beats are then normalized in time to be with the fixed length (120) samples where the beat spans on  $BI \times 120/1.5$  samples. The rest of the 120 samples are filled with zeros.

#### 4.2.6 DL-Based BP reconstruction from rPPG estimated signals

The estimated rPPG signal is segmented into beats and the time interval is recorded. Beat selection is applied to the segmented rPPG beats and then these beats are normalized. The normalized selected beats are used along with the beat intervals for BP prediction using the trained network described in Subsection 4.1.7.

## 5 Experimental results

The five trained models are evaluated based on the contact PPG dataset and real videos. Also, the performance is evaluated under the original labeling and the logarithmic labeling.

### 5.1 Segmentation results

We performed an ablation study on the modified SegNet model to gain a thorough understanding of the effects of replacing the VGG encoder network with the ResNet encoder network. The segmentation results for the skin dataset proposed on [130] shown in Table 1 show that the modification significantly improved the segmentation in terms of accuracy metric by 6% and pixel-level dice coefficient metric by 8%.

**Table 1** Evaluation of SegNet skin segmentation model with VGG and ResNet encoders

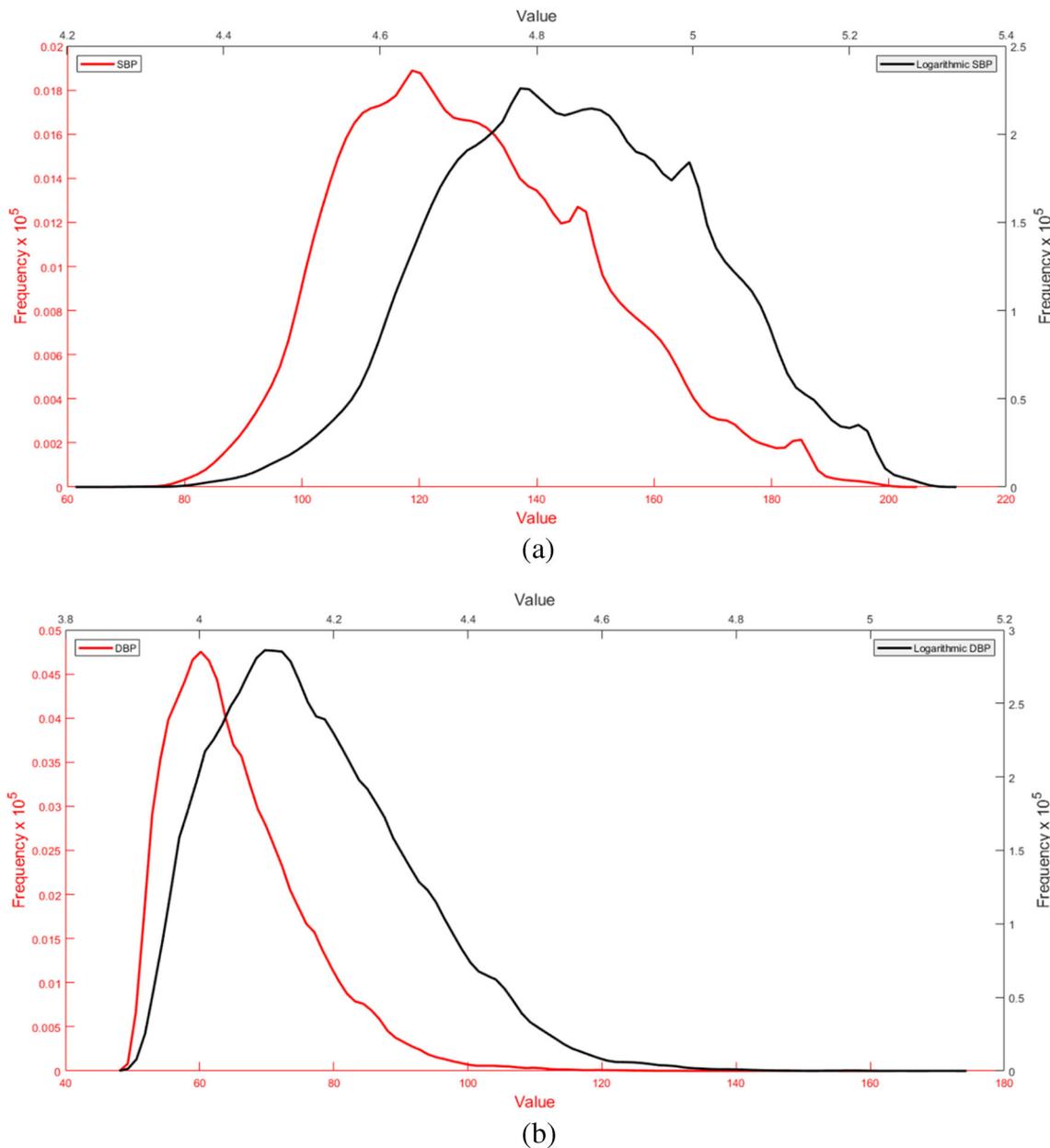
Model	Accuracy	Dice coefficient
SegNet (VGG encoder)	91.50 %	87.24 %
SegNet (ResNet encoder)	97.02%	95.43%

## 5.2 Contact dataset and the logarithmic transformation

Thanks to the Physionet's MIMIC II dataset (Multi-parameter Intelligent Monitoring in Intensive Care) [42–45] that provides joint PPG-ABP data needed for feeding the learning models. A compiled version of that dataset is introduced by [117] where it has a better presentation. However, by inspecting that dataset, 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-based

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 signal and corresponding ABP as a label. For proper handling and filtering, records are segmented into sections of 1000 samples in length. So, we have 30,660 records. These signals are segmented into 309860 beats. Data cleaning results in 140538 beats.

The distributions of the SBP and DBP of the data are shown in Fig. 6, based on logarithmic and original labeling.



**Fig. 6** Logarithmic labeling versus original labeling distribution (a) SBP Distribution. (b) DBP Distribution

It is clear that the logarithmic labeling shifts the distribution partially to the higher range.

### 5.2.1 BP estimation from cPPG signals

The dataset was split into training, validation, and test sets on a beat basis to prevent contamination of the validation and test set by training data. We used 130690 beats for training and 9848 beats for validation and testing. Input pipelines and NNs were implemented using TensorFlow 2.4 and Python 3.9 was used for training (Adam optimizer,  $\alpha = 0.001$ , MSE loss, 200 epochs). We used the models with the lowest MAE on the validation set for further testing.

We used the MAE and standard deviation metrics to assess the performance of all methods. Different networks are implemented and tested for each SBP and DBP estimation. The prediction errors for the test dataset are determined.

Table 2 shows the evaluation results of different DL networks; including AlexNet [90, 131], ResNet [90, 129], Slapničar [83, 90] and Mean Regression [132] in comparison with the proposed system with five different NNs adopted including ResNet, VGG16, AlexNet, DenseNet and MobileNet. The evaluation is based on the standard deviations and the mean absolute errors for each of the SBP errors and DBP errors. Also, Table 2 summarises the method and the main idea of each system as well. From Table 2, it can be recognized that using the proposed system including pre-processing and network training with 2D images outperforms the state-of-the-art systems; AlexNet, ResNet, Slapničar, and Mean Regression in the sense of MAE and errors' standard deviation. The enhancement of the proposed system is due to four forks which are:

- Enhancement in the ROI selection by using accurate skin segmentation.
- Signal segmentation into beats that gives the flexibility to reject individual beats.
- Data cleaning before training that is by rejecting invalid signals and invalid beats.
- Beat-by-beat training along with beat interval embedding implicitly

### 5.2.2 Effect of using logarithmic labels on learning

Logarithmic transform is a useful pre-processing technique, especially for skewed distributions. Log-transformation is used when the data is highly skewed. It makes the distributions more aligned towards the normal distribution curve for better learning and thus prediction. The Log transformation method can significantly improve accuracy [133]. Logarithmic transforms can help convert skewed distributions into more normal distributions. Logarithmic transformation compresses large values closer to smaller values, reducing the range of values. This can help algorithms handle features that vary a lot in scale. Improves gradient descent convergence. By reducing large values, logarithmic transforms can speed up the convergence of gradient descent algorithms used in training. The gradients don't change as drastically with each update.

Table 3 shows a comparison between using SBP and DBP values as labels and using logarithmic SBP and DBP as labels. From this table, it can be shown that using logarithmic labels with resNet101 gives better results compared to using BP values as labels in the case of DBP. This is because using logarithmic labels overcomes the distribution skewness that exists

**Table 2** Evaluation results of the DL Networks

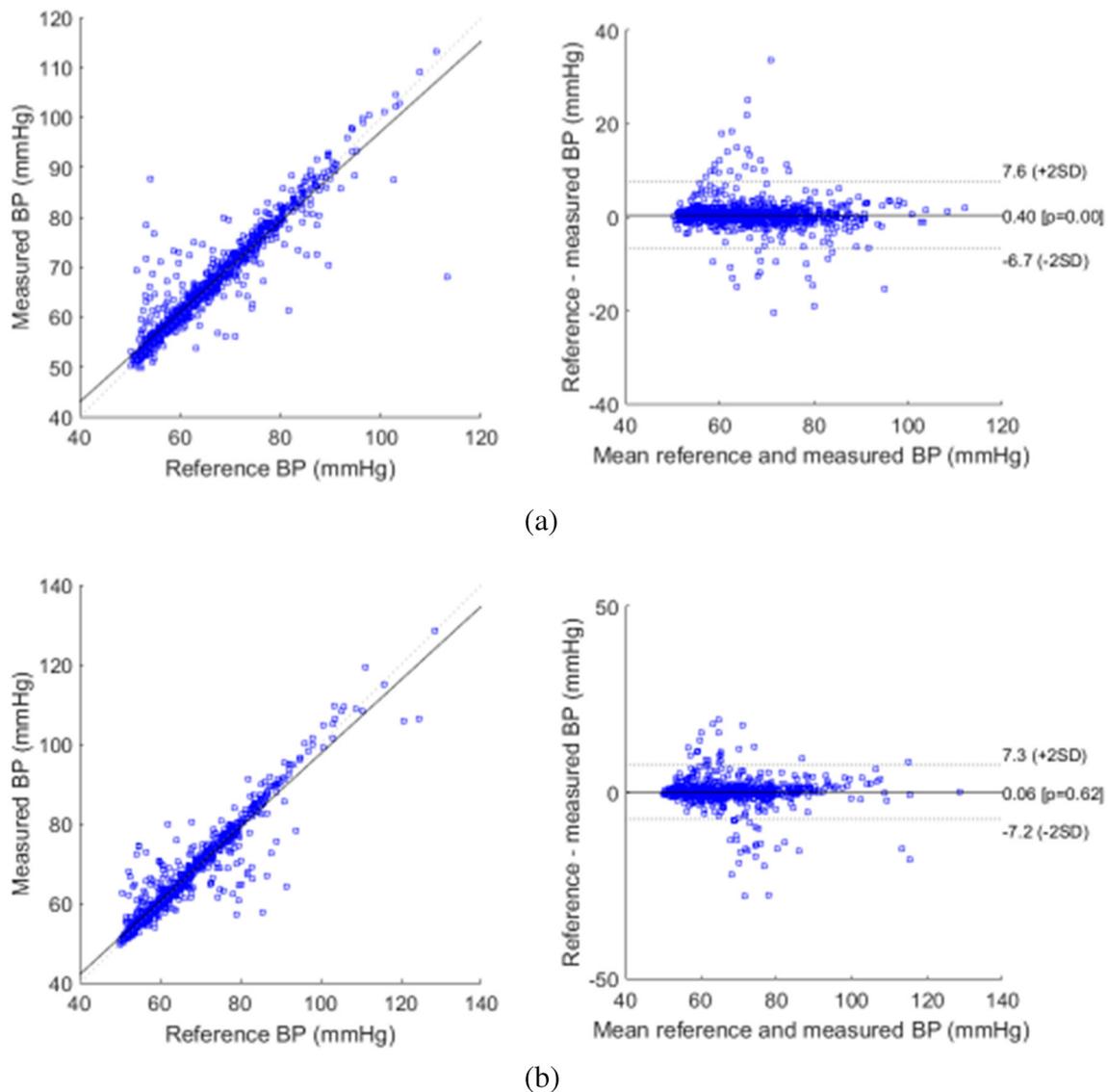
DL-Net	STD (SBP)	Mean (SBP)	STD (DBP)	Mean (DBP)	Data Used	Method Used
AlexNet [90, 131]	9.11	15.20	4.92	8.52	MIMIC III	Using AlexNet, PPG time series as inputs
ResNet [90, 129]	12.61	12.51	9.84	8.3	MIMIC III	Using ResNet, PPG time series as inputs
Slapničar [83, 90]	10.12	13.56	5.02	8.5	MIMIC III	Temporal and frequency features of PPG, PPG' and PPG'', deep learning (spectro-temporal ResNet)
Mean Regression [132]	10.4	13.9	5.0	8.5	MIMIC III	Use ML regression
2D VGG16 with proposed system	5.8285	2.7155	4.1021	1.9400	MIMIC II	Using 2D VGG16, PPG and BI
2D AlexNet with proposed system	6.5727	3.8795	4.3613	2.3584	MIMIC II	Using 2D AlexNet with Logarithmic labels, PPG and BI
2D DenseNet with proposed system	6.8502	4.0735	4.6267	2.6067	MIMIC II	Using 2D DenseNet with Logarithmic labels, PPG and BI
2D MobileNet with proposed system	8.3671	5.3786	5.1208	3.1796	MIMIC II	Using 2D MobileNet with Logarithmic labels, PPG and BI
2D ResNet with proposed system	5.4782	2.3453	3.8539	1.6854	MIMIC II	Using 2D ResNet101 with Logarithmic labels, PPG and BI

**Table 3** The effect of using the logarithmic labels on two different DL networks

DL-Net	STD (SBP)	Mean (SBP)	STD (DBP)	Mean (DBP)
2D VGG16 without logarithmic	5.57	2.62	4.02	2.05
2D VGG16 with logarithmic	5.8285	2.7155	4.0121	1.9400
2D Resnet101 without logarithmic	5.44	2.36	3.90	1.80
2D Resnet101 with logarithmic	5.4782	2.3453	3.8539	1.6854

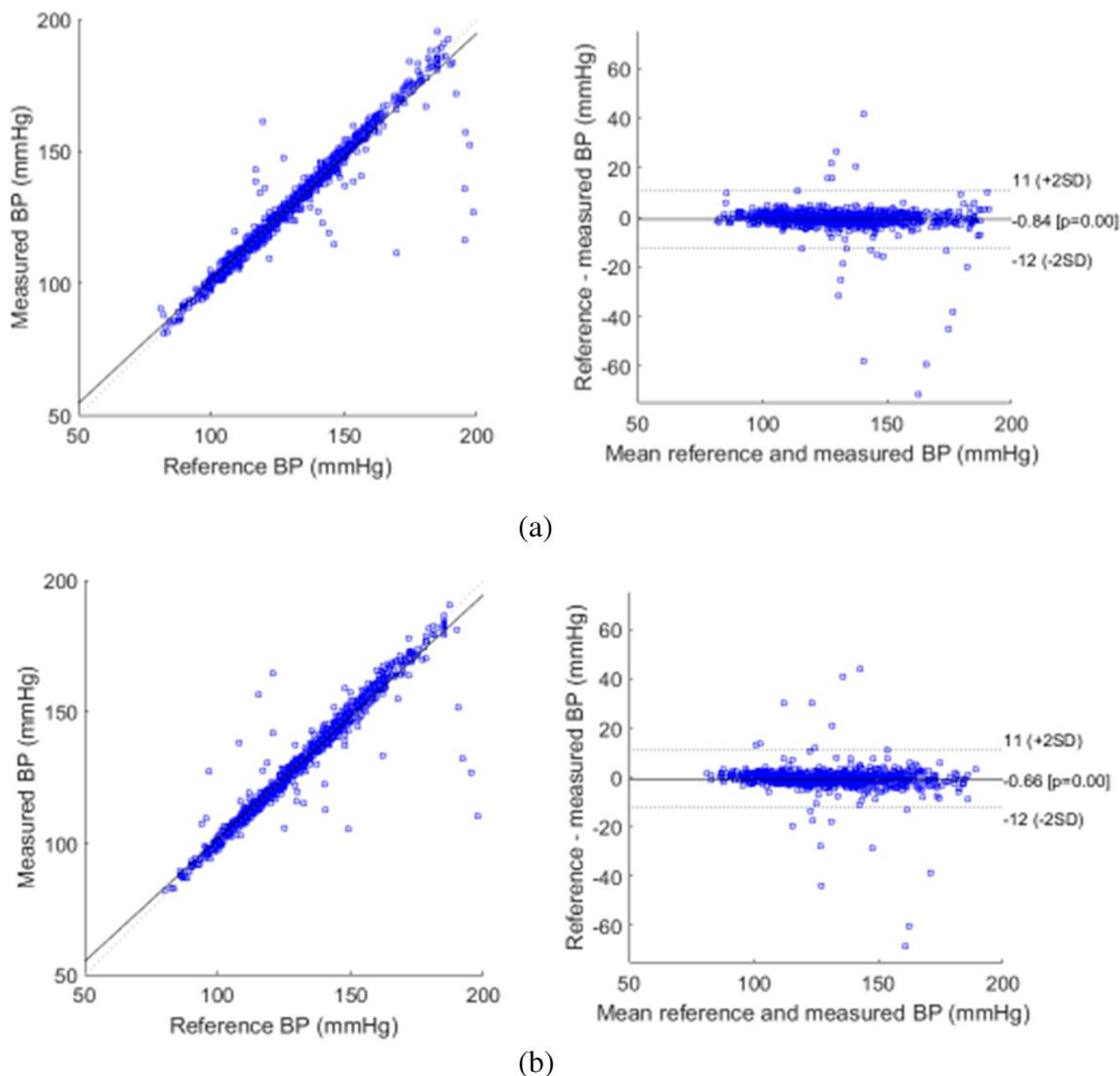
in the DBP data. However, using logarithmic SBP labels has a very small effect on the prediction results as the SBP distribution is not skewed. To evaluate the predicted SBP and DBP, Bland-Altman [134] plots are used. Bland-Altman plots are

extensively used to evaluate the agreement among predicted and ground truth values for SBP and DBP. Bland-Altman plots allow the identification of any systematic difference between the measurements or possible outliers. Figures 7



**Fig. 7** (a) The Bland-Altman plot for the estimated DBP values from cPPG using ResNet101-based 2D beat-by-beat network: (left) correlation between the estimated DBP and the ground truth DBP values, (right) the DBP error vs. the mean DBP error. (b) The Bland-Altman plot for

the estimated DBP values from cPPG using ResNet101-based 2D beat-by-beat network with logarithmic labels: (left) correlation between the estimated DBP and the ground truth DBP values, (right) the DBP error vs. the mean DBP error



**Fig. 8** (a) The Bland-Altman plot for the estimated SBP values from CPPG using ResNet101-based 2D beat-by-beat network: (left) correlation between the estimated SBP and the ground truth SBP values, (right) the SBP error vs. the mean SBP error. (b) The Bland-Altman plot for

the estimated SBP values from CPPG using ResNet101-based 2D beat-by-beat network with logarithmic labels: (left) correlation between the estimated SBP and the ground truth SBP values, (right) the SBP error vs. the mean SBP error

and 8 show the Bland-Altman plots for DBP and SBP using the ResNet101 network. From these figures, we can recognize the high correlation between the estimated DBP/SBP and the corresponding ground truth.

The main goal of using a logarithmic transformation is to modify the skewed data to become more suitable for analysis. Since the diastolic data is clearly skewed as shown in Fig. 6b, logarithmic transformation is more effective in case of diastolic data compared to systolic data as depicted in Table 3. The slight effect of logarithmic transformation in case of systolic is due to slight skewness in the original data. However, the skewness in diastolic data is higher, therefore the logarithmic transformation modifies the skewness and thus the

mean absolute error is decreased by using the logarithmic transformation.

To ensure that the accuracy improvement is due to the logarithmic transformation, not training variability, all evaluated

**Table 4** Correlation coefficients for different DL Networks

DL-Net	CC (SBP)	CC (DBP)
2D VGG16	0.9381	0.9539
2D Resnet101	0.9434	0.9644
2D VGG16 with logarithmic	0.938	0.9644
2D Resnet101 with logarithmic	0.9487	0.9644

**Table 5** Evaluation results of the DL Networks with cPPG data

DL-Net	SBP		DBP	
	STD	Mean	STD	Mean
Proposed BbB-based CNN	13.8582	10.1383	7.7160	5.4848
Signal-based [94]	21.3832	17.5138	10.1268	7.4355

models were trained and evaluated under the same training conditions. To achieve this, employing pretrained models for the training process in both scenarios (without/with logarithmic transformation) guarantees consistent initial weights and eliminates randomness as a factor. Moreover, we used the same train-test dataset for all models, the same loss function, optimizer, and hyperparameters- including learning rate, number of epochs, and patch size. This meticulous approach helps eliminate variations in the training process, ensuring a fair and reliable comparison of the effects of the logarithmic transformation on accuracy improvements which is consistent with the conclusion reported by the authors in [133].

The correlation coefficients (CC) between the predicted and the ground truth values are tabulated in Table 4. This table confirms the results of the Bland-Altman figure that using logarithmic labels increases the correlation between the predicted and ground truth values and that ResNet101 results in a higher correlation than VGG16 networks.

### 5.2.3 Effect of per beat learning

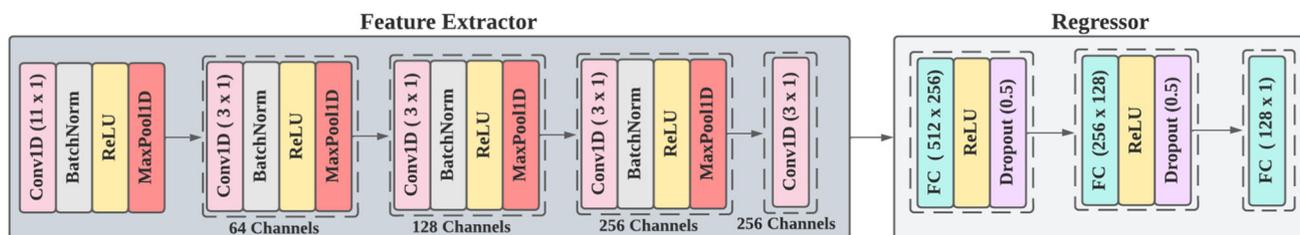
Table 5 shows the evaluation results of the two approaches; the signal-based DL approach and the beat-by-beat DL approach. The evaluation is based on the standard deviation and the mean absolute error for each of the SBP errors and DBP errors. From this table, it can be shown that the proposed approach achieves an enhancement in terms of MAE of 7.36 and 1.95 for SBP and DBP, respectively. Also, in terms of standard deviation, the proposed approach achieves enhancements of 7.5 and 2.41 for SBP and DBP, respectively. The enhancement of the proposed approach is due to two reasons; the first is that the training is based on beat-by-beat and thus data cleaning in beat level is more accurate and invalid beats are rejected rather than rejecting the whole

signal. Also, because beat-by-beat prediction allows for the rejection of invalid beats, the predicted BP that corresponds to the valid beats is averaged. The second is that the signal-based approach is based on the estimation of the ABP signal from the PPG signal, and then the SBP is calculated as the maximum value of the predicted ABP and the DBP is the minimum value of the predicted ABP. The calculation of the maximum and minimum values over the whole signal, which may contain invalid beats, leads to inaccurate values.

### 5.3 2D-DL network versus 1D-DL network

As mentioned in Section 4.1.7, 1D beats are converted to 2D images to get the benefits of using pre-trained networks. To show the benefits of using 2D images rather than 1D signals, we have evaluated 1D Beat-by-beat CNN network with BP estimation in comparison to 2D Beat-by-beat CNN network.

The 1D DL network is composed of two main components, namely, the feature extractor and the regressor. The feature extractor is a 5-layer convolutional neural network, while the regressor is a 3-layer fully connected. The detailed architecture of the feature extractor and the regressor is further illustrated in Fig. 9. In particular, the first layer is composed of a 1-D convolutional filter with a kernel size of 11x1 and 32 channels. Then, a batch normalization layer is used to improve the performance and increase the training speed. In this model, we used the rectified linear unit (ReLU) as a non-linear activation function. The subsequent 4 layers have a kernel size of 3 x 1 with an increasing number of channels that reaches 256 for the last layer. On the other hand, the regressor network is mainly composed of 3 fully connected (FC) layers with output dimensions of 256, 128, and 1 respectively. Each FC layer is followed by ReLU as non-linear activation and a dropout layer as a regularization technique.



**Fig. 9** The 1D deep-learning network architecture

**Table 6** Evaluation results of 1D and 2D DL Networks with cPPG data

DL-Net	SBP		DBP	
	STD	Mean	STD	Mean
Proposed BbB-based CNN	13.8582	10.1383	7.7160	5.4848
1D BbB-based CNN	16.8943	16.9942	8.7254	6.3644

The evaluation results are depicted in Table 6. From this table, it can be shown that using 2D DL network outperforms the 1D DL network for SBP and DBP estimation in the sense of mean and standard deviation of errors. One of the main advantages of the 2D network is the utilization of a pre-trained network using transfer learning, which enables the 2D network to outperform the 1D network.

#### 5.4 BP estimation from estimated rPPG signals

To evaluate the proposed system with the real videos, we conducted 19 experiments that are by recording 19 digital videos 1 minute in length each. We use a digital camera with  $1980 \times 1080$  spatial resolution and 60 frames per second (FPS). The distance between the subject and the camera is about 1m. The subject is asked to sit in front of the camera and not moving as possible. The room light is used as

the source of light. The subjects are all members of our lab. Simultaneously, an OMRON M2 1030 is used to measure the BP to be used as reference values. Each video is subjected to the prediction phase of the proposed system. Tables 7 and 8 show the results of different DL networks. This table shows the SBP and DBP for 19 subjects including white and black skins. The results tabulated in Tables 7 and 8 are obtained from the first 10 seconds of each video. Each video goes through the prediction phase of the proposed system shown in Fig. 2. As the proposed system works on a beat-by-beat basis, the resulting BP for the selected beats is averaged to get one value for each SBP and DBP within the prediction period (10 seconds) that is to reduce the processing time and for a fair comparison with the Omron M2 readings as it takes about 15 seconds to get a reading. Actually, it is difficult to compare the proposed system with that of the state-of-the-art systems in the case of real videos. This is due to many

**Table 7** Evaluation results of the DL Networks for DBP using real videos

vid #	GT DBP	2D Resnet101	2D Resnet101 with Logarithmic
1	68	68.593898	69.683993
2	79	64.459155	67.216491
3	75	68.602373	71.431188
4	82	69.513304	72.003017
5	73	61.047254	66.094239
6	71	72.809106	73.132951
7	65	65.056057	71.015176
8	77	64.349126	68.504944
9	58	62.188442	62.350533
10	52	69.605949	60.083639
11	68	57.731594	58.703457
12	68	58.409169	61.025874
13	71	69.054068	73.952834
14	69	68.247353	71.65051
15	69	61.900433	65.608894
16	77	66.277325	69.241883
17	69	61.584206	70.032854
18	83	73.645467	76.247597
19	72	65.007325	69.12037
	MAE	7.706595	5.616028
	STD	7.804062	6.013052

**Table 8** Evaluation results of the DL Networks for SBP using real videos

vid #	GT SBP	2D Resnet101	2D Resnet101 with Logarithmic
1	119	123.02007	119.31379
2	127	128.22113	128.0279
3	125	119.5899	111.13243
4	127	120.1167	113.46644
5	109	121.10866	119.51879
6	113	111.69385	120.14906
7	109	115.2344	114.12664
8	109	114.48332	120.24121
9	103	107.76915	108.31242
10	99	122.12389	110.45409
11	115	127.40166	132.87297
12	110	123.25407	115.87918
13	121	114.56446	116.07251
14	118	122.48306	122.04311
15	119	133.19645	102.29696
16	121	120.2322	118.94467
17	132	133.22771	116.64305
18	125	130.90062	115.00311
19	119	114.22384	109.26411
	MAE	7.052803	8.742941
	STD	8.027882	10.42648

factors including the privacy factor where it is not allowed to share real videos. Also, preprocessing steps are not always available with shared networks' weights. Moreover, even the shared DL codes refer to the public database without performing the preprocessing steps. Therefore, we evaluate the proposed system with our real videos for the study of different scenarios. Tables 7 and 8 reveal that using logarithmic labels gives better results (lowest STD and lowest MAE) than that using normalized values as labels in the case of SBP. While in the case of DBP, using normalized labels gives better results (lowest STD and lowest MAE) than that using normalized values as logarithmic labels.

## 5.5 Complexity evaluation

In the context of deep learning, the complexity of a model is often measured by the number of parameters it has and the float point operations per second (FLOPs) required to train and deploy the model. In general, models with a large number of parameters require more data to train, and they can be computationally expensive to deploy. FLOPs simply mean the total floating point operations required for a single forward pass. The higher the FLOPs, the slower the model and hence low throughput.

Table 9 shows the number of parameters and FLOPS for the deep learning models that were proposed. The number of parameters is a measure of the model's complexity, and the FLOPS is a measure of the model's computational cost. As you can see, the number of parameters and FLOPS can vary widely between different deep learning models. In the case of the Resnet101 and VGG16 models, despite having more parameters (< 42 million and < 39 million, respectively) and requiring more FLOPS to train (< 7 billion and < 15 billion, respectively), they achieved the best BP estimation results. These results suggest that the Resnet101 and VGG16 models are more complex than the other models, but they are also more accurate. This is likely because the additional complexity allows these models to learn more complex relationships between the input and output data.

**Table 9** Complexity evaluation for the used DL Networks

Model	No. of Parameters (In Millions)	FLOPs (In Millions)
2D VGG16	39.81	$15.38 \times 10^3$
2D AlexNet	3.39	656.49
2D DensNet	7.06	$2.87 \times 10^3$
2D MobileNet	2.35	313.05
2D Resnet101	42.71	$7.84 \times 10^3$

## 5.6 Uncertainty and explainability

For future work and to build a reliable system, it is important to study addition factors to overcome uncertainty situation. Uncertainty can be caused by a number of factors, including a mismatch between training and testing data and variances in the data acquisition systems [135, 136]. In our system, the change in light and the skin color may need more investigations. Furthermore, to understand how the Artificial Intelligence (AI) model makes a particular prediction, Explainable Artificial Intelligence (XAI) methods such as SHapley Additive exPlanation (SHAP) and Locally Interpretable Model Agnostic Explanations (LIME) are demanded]. XAI provides insights into AI models to aid in the enhancement of their reliability, robustness, and performance [135–138].

## 6 Conclusion

In this paper, we presented a beat-to-beat BP estimation system from facial videos. To cope with the challenge of patients' motion as a source of error, face tracking, and skin segmentation are employed. The training PPG data is cleaned strictly. Transfer learning is applied based on the well-trained image deep learning networks. The 1D beat is mapped to a 2D image. The cleaning metrics are applied to the training data and the extracted rPPG signal as well. Only the selected (valid) beats are applied to the trained DL network to predict BP. Also, the monitoring rate increases where the probability of catching individually non-distorted beats is much higher than the probability of catching a complete regular signal (multiple beats in a successive stream). Skewness in data distribution is adapted partially in the high BP range through logarithmic transformation. Five deep learning networks are tested including AlexNet, Resnet101, VGG16, MobileNet, and DenseNet. Resnet101 network outperforms other tested networks. Also, logarithmic labeling outperforms the original labeling. Based on the experimental results, the proposed system outperforms the state-of-the-art systems in the sense of MAE and standard deviation.

The main limitation of the proposed system is the computational load and that it is still difficult to work in real-time continuous BP monitoring. Therefore, in our current and future work, we are focusing on how to reduce the computational time of the proposed system so that it can be used in continuous real-time BP monitoring.

**Funding** Open access funding provided by The Science, Technology & Innovation Funding Authority (STDF) in cooperation with The Egyptian Knowledge Bank (EKB). This work is supported by the Information Technology Industry Development Agency (ITIDA) under number ARP2019.R27.4

**Data Availability** In this work, we utilized the MIMIC II dataset which is a public dataset. It is available at: <https://peterhcharlton.github.io/RRest/mimiciidataset.html>. (Accessed on February 18, 2024) Our data will be available upon reasonable request.

## Declarations

**Conflict of Interest** The authors declare that there is no Conflict of Interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Organization, WH (2022) Cardiovascular diseases, [https://www.who.int/health-topics/cardiovascular-diseases#tab=tab\\_1](https://www.who.int/health-topics/cardiovascular-diseases#tab=tab_1), (Accessed on september, 22, 2022)
- Lam S, et al (2021) Intraoperative Invasive Blood Pressure Monitoring and the Potential Pitfalls of Invasively Measured Systolic Blood Pressure. *Cureus* 13(8)
- Meidert AS, Saugel B (2018) Techniques for non-invasive monitoring of arterial blood pressure. *Front Med* 4:231
- Geddes LA (2013) Handbook of blood pressure measurement. Springer Science & Business Media
- American Society of Anesthesiologists. Standards of the American Society of Anesthesiologists: Standards for Basic Anesthetic Monitoring; Available from: <https://www.asahq.org/standards-and-guidelines/standards-for-basic-anesthetic-monitoring> Accessed on February, 2022. (2020)
- Liao J, Liu D, Su G, Liu L (2021) Recognizing diseases with multivariate physiological signals by a DeepCNN-LSTM network. *Appl Intell* 1–13
- Xu W, Huang M-C (2015) Total health: Toward continuous personal monitoring, in *Wearable Electronics Sensors*. Springer. pp 37–56
- McGillion MH, Allan K, Ross-Howe S et al (2022) Beyond wellness monitoring: continuous multiparameter remote automated monitoring of patients. *Canadian J Cardiol* 38(2):267–278
- Le T et al (2020) Continuous non-invasive blood pressure monitoring: a methodological review on measurement techniques. *IEEE Access* 8:212478–212498
- Panula T et al (2022) Advances in non-invasive blood pressure measurement techniques. *IEEE Rev Biomed Eng*
- Mejia-Mejia E et al (2022) Photoplethysmography signal processing and synthesis. In: *Photoplethysmography*, Elsevier. pp 69–146
- Loh HW et al (2022) Application of photoplethysmography signals for healthcare systems: An in-depth review. *Comput Methods Programs Biomed* 216:106677
- Almarshad MA et al (2022) Diagnostic features and potential applications of PPG signal in healthcare: A systematic review. in *Healthcare*. MDPI
- Singstad B-J et al (2021) Estimation of heart rate variability from finger photoplethysmography during rest, mild exercise and mild mental stress. *J Electrical Bioimpedance* 12(1):89–102
- Kuwalek P et al (2021) Research on methods for detecting respiratory rate from photoplethysmographic signal. *Biomed Signal Process Control* 66:102483
- Martínez G et al (2018) Can photoplethysmography replace arterial blood pressure in the assessment of blood pressure? *J Clinical Med* 7(10):316
- Elgendi M et al (2019) The use of photoplethysmography for assessing hypertension. *NPJ Digital Med* 2(1):1–11
- Omer OA et al (2022) Beat-by-Beat ECG Monitoring from Photoplethysmography Based on Scattering Wavelet Transform. *Traitement du Signal* 39(5)
- Zhu Q et al (2021) Learning your heart actions from pulse: ECG waveform reconstruction from PPG. *IEEE Int Things J* 8(23):16734–16748
- Abdelgaber KM et al (2023) Subject-Independent per Beat PPG to Single-Lead ECG Mapping. *Information* 14(7):377
- Kavsaoğlu AR, Polat K, Hariharan M (2015) Non-invasive prediction of hemoglobin level using machine learning techniques with the PPG signal's characteristics features. *Appl Soft Comput* 37:983–991
- Haque CA et al (2021) Comparison of Different Methods to Estimate Blood Oxygen Saturation using PPG. In: 2021 International conference on information and communication technology convergence (ICTC). IEEE
- Liu Z-D et al (2023) Cuffless Blood Pressure Measurement using Smartwatches: A Large-scale Validation Study. *IEEE J Biomed Health Inform*
- Zhao L et al (2023) Emerging sensing and modeling technologies for wearable and cuffless blood pressure monitoring. *NPJ Digital Med* 6(1):93
- Ray I et al (2021) Skin tone, confidence, and data quality of heart rate sensing in WearOS smartwatches. In: 2021 IEEE International conference on pervasive computing and communications workshops and other affiliated events (PerCom Workshops). IEEE
- Qureshi F, Krishnan S (2018) Wearable hardware design for the internet of medical things (IoMT). *Sensors* 18(11):3812
- Zauneder S et al (2018) Cardiovascular assessment by imaging photoplethysmography—a review. *Biomed Eng/Biomedizinische Technik* 63(5):617–634
- Kumar M, Veeraraghavan A, Sabharwal A (2015) DistancePPG: Robust non-contact vital signs monitoring using a camera. *Biomed Optics Express* 6(5):1565–1588
- Hassan MA et al (2017) Heart rate estimation using facial video: A review. *Biomed Signal Process Control* 38:346–360
- Moço AV, Stuijk S, de Haan G (2018) New insights into the origin of remote PPG signals in visible light and infrared. *Scientific Report* 8(1):1–15
- Kamshilin AA et al (2015) A new look at the essence of the imaging photoplethysmography. *Scientific Reports* 5(1):1–9
- Zhou Y et al (2019) The noninvasive blood pressure measurement based on facial images processing. *IEEE Sensors J* 19(22):10624–10634
- McDuff D (2023) Camera measurement of physiological vital signs. *ACM Comput Surv* 55(9):1–40
- Matsumura K et al (2018) Cuffless blood pressure estimation using only a smartphone. *Sci Reports* 8(1):1–9
- Luo H et al (2019) Smartphone-based blood pressure measurement using transdermal optical imaging technology. *Circulation: Cardiovascular Imaging* 12(8):e008857

36. Qayyum A et al (2022) Assessment of physiological states from contactless face video: a sparse representation approach. *Computing* 1–21
37. Gudi A, Bittner M, van Gemert Jv, (2020) Real-Time Webcam Heart-Rate and Variability Estimation with Clean Ground Truth for Evaluation. *Appl Sci* 10(23):8630
38. Salah M et al (2022) Robust Facial-Based Inter-Beat Interval Estimation Through Spectral Signature Tracking and Periodic Filtering. In: *Intelligent Sustainable Systems*. Singapore: Springer Singapore
39. Steinman J et al (2021) Smartphones and Video Cameras: Future Methods for Blood Pressure Measurement. *Front Digital Health* 3
40. Hosni A, Atef M (2023) Remote real-time heart rate monitoring with recursive motion artifact removal using PPG signals from a smartphone camera. *Multimed Tools Appl* 1–18
41. Wang W et al (2016) Algorithmic principles of remote PPG. *IEEE Trans Biomed Eng* 64(7):1479–1491
42. Saeed M et al (2002) MIMIC II: a massive temporal ICU patient database to support research in intelligent patient monitoring. in *Computers in cardiology*. IEEE
43. Johnson AE et al (2016) MIMIC-III, a freely accessible critical care database. *Scientific Data* 3(1):1–9
44. Liang Y et al (2018) A new, short-recorded photoplethysmogram dataset for blood pressure monitoring in China. *Sci Data* 5(1):1–7
45. Goldberger AL et al (2000) PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation* 101(23):e215–e220
46. Qin K et al (2022) Machine learning and deep learning for blood pressure prediction: a methodological review from multiple perspectives. *Artif Intell Rev* 1–102
47. Yin R-N et al (2021) Heart rate estimation based on face video under unstable illumination. *Appl Intell* 1–17
48. Rahman H, Ahmed MU, Begum S (2019) Non-contact physiological parameters extraction using facial video considering illumination, motion, movement and vibration. *IEEE Trans Biomed Eng* 67(1):88–98
49. Mironenko Y et al (2020) Remote photoplethysmography: Rarely considered factors. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*
50. Verkruysse W, Svaasand LO, Nelson JS (2008) Remote plethysmographic imaging using ambient light. *Optics Express* 16(26):21434–21445
51. Man P-K et al (2022) Blood Pressure Measurement: From Cuff-Based to Contactless Monitoring. in *Healthcare MDPI*
52. González S, Hsieh W-T, Chen TP-C (2023) A benchmark for machine-learning based non-invasive blood pressure estimation using photoplethysmogram. *Scientific Data* 10(1):149
53. Schrupf F et al (2021) Assessment of Non-Invasive Blood Pressure Prediction from PPG and rPPG Signals Using Deep Learning. *Sensors* 21(18):6022
54. Boccignone G et al (2020) An Open Framework for Remote-PPG Methods and their Assessment. *IEEE Access* 8:216083–216103
55. Nour M et al (2023) A Novel Cuffless Blood Pressure Prediction: Uncovering New Features and New Hybrid ML Models. *Diagnostics* 13(7):1278
56. Bulhões da Silva Costa T et al (2023) Blood pressure estimation from photoplethysmography by considering intra-and inter-subject variabilities: guidelines for a fair assessment
57. Liu J et al (2023) A novel interpretable feature set optimization method in blood pressure estimation using photoplethysmography signals. *Biomed Signal Process Control* 86:105184
58. Mukhlif AA, Al-Khateeb B, Mohammed MA (2022) An extensive review of state-of-the-art transfer learning techniques used in medical imaging: Open issues and challenges. *J Intell Syst* 31(1):1085–1111
59. Magdalena Nowara E et al (2018) SparsePPG: Towards driver monitoring using camera-based vital signs estimation in near-infrared. In: *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*
60. Xi L et al (2022) Weighted Combination and Singular Spectrum Analysis Based Remote Photoplethysmography Pulse Extraction in Low-light Environments. *Med Eng Phys* 103822
61. Zhao C et al (2018) A novel framework for remote photoplethysmography pulse extraction on compressed videos. In: *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*
62. Wang W et al (2017) Robust heart rate from fitness videos. *Physiological Measure* 38(6):1023
63. Lovisotto G et al (2020) Seeing red: PPG biometrics using smartphone cameras. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*
64. Bousefsaf F et al (2021) iPPG 2 cPPG: Reconstructing contact from imaging photoplethysmographic signals using U-Net architectures. *Comput Biol Med* 138:104860
65. Wang D et al (2020) Detail-preserving pulse wave extraction from facial videos using consume-level camera. *Biomed Optics Express* 11(4):1876–1891
66. Pham C et al (2021) Effectiveness of consumer-grade contactless vital signs monitors: a systematic review and meta-analysis. *J Clinical Monitoring Comput* 1–14
67. Fouad R, Omer OA, Aly MH (2019) Optimizing remote photoplethysmography using adaptive skin segmentation for real-time heart rate monitoring. *IEEE Access* 7:76513–76528
68. Bobbia S et al (2019) Unsupervised skin tissue segmentation for remote photoplethysmography. *Pattern Recognition Lett* 124:82–90
69. Pirnar Ž, Finžgar M, Podražaj P (2021) Performance Evaluation of rPPG Approaches with and without the Region-of-Interest Localization Step. *Appl Sci* 11(8):3467
70. Zaunseder S et al (2022) Signal-to-noise ratio is more important than sampling rate in beat-to-beat interval estimation from optical sensors. *Biomed Signal Process Control* 74:103538
71. Wander J, Morris D (2014) A combined segmenting and non-segmenting approach to signal quality estimation for ambulatory photoplethysmography. *Physiological Measurement* 35(12):2543
72. Fischer C et al (2017) Extended algorithm for real-time pulse waveform segmentation and artifact detection in photoplethysmograms. *Somnologie* 21(2):110–120
73. Elgendi M et al (2013) Systolic peak detection in acceleration photoplethysmograms measured from emergency responders in tropical conditions. *PLoS One* 8(10):e76585
74. Li P et al (2020) Video-Based Pulse Rate Variability Measurement Using Periodic Variance Maximization and Adaptive Two-Window Peak Detection. *Sensors* 20(10):2752
75. Kachuee M et al (2016) Cuffless blood pressure estimation algorithms for continuous health-care monitoring. *IEEE Trans Biomed Eng* 64(4):859–869
76. Ding X, Zhang Y-T (2019) Pulse transit time technique for cuffless unobtrusive blood pressure measurement: from theory to algorithm. *Biomed Eng Lett* 9(1):37–52
77. Liang Y, Abbott D, Howard N, Lim K, Ward R, Elgendi M (2019) How effective is pulse arrival time for evaluating blood pressure? Challenges and recommendations from a study using the MIMIC database. *J Clinical Med*
78. Clifford GD, Scott DJ, Villarreal M (2010) User guide and documentation for the MIMIC II database (version 2, release 1)
79. Finnegan E et al (2023) Features from the photoplethysmogram and the electrocardiogram for estimating changes in blood pressure. *Scientific Reports* 13(1):986

80. Khalid SG et al (2018) Blood pressure estimation using photoplethysmography only: comparison between different machine learning approaches. *J Healthcare Eng* 2018
81. Slapničar G, Luštrek M, Marinko M (2018) Continuous blood pressure estimation from PPG signal. *Informatica* 42(1)
82. Yan W-R et al (2019) Cuffless continuous blood pressure estimation from pulse morphology of photoplethysmograms. *IEEE Access* 7:141970–141977
83. Slapničar G, Mlakar N, Luštrek M (2019) Blood pressure estimation from photoplethysmogram using a spectro-temporal deep neural network. *Sensors* 19(15):3420
84. Xing X, Sun M (2016) Optical blood pressure estimation with photoplethysmography and FFT-based neural networks. *Biomed Optics Express* 7(8):3007–3020
85. Liang Y et al (2018) Photoplethysmography and deep learning: enhancing hypertension risk stratification. *Biosensors* 8(4):101
86. Wang W et al (2023) PulseDB: A large, cleaned dataset based on MIMIC-III and VitalDB for benchmarking cuff-less blood pressure estimation methods. *Front Digital Health* 4:1090854
87. Salah M et al (2022) Beat-Based PPG-ABP Cleaning Technique for Blood Pressure Estimation. *IEEE Access* 10:55616–55626
88. Schrupf F, Serdack PR, Fuchs M (2022) Regression or Classification? Reflection on BP prediction from PPG data using Deep Neural Networks in the scope of practical applications. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*
89. Ali NF, Atef M (2023) An efficient hybrid LSTM-ANN joint classification-regression model for PPG based blood pressure monitoring. *Biomed Signal Process Control* 84:104782
90. Schrupf F et al (2021) Assessment of deep learning based blood pressure prediction from PPG and rPPG signals. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*
91. Hsu Y-C et al (2020) Generalized deep neural network model for cuffless blood pressure estimation with photoplethysmogram signal only. *Sensors* 20(19):5668
92. Qin C et al (2023) Cuff-Less Blood Pressure Prediction Based on Photoplethysmography and Modified ResNet. *Bioengineering* 10(4):400
93. Kanoga S et al (2023) Comparison of seven shallow and deep regressors in continuous blood pressure and heart rate estimation using single-channel photoplethysmograms under three evaluation cases. *Biomed Signal Process Control* 85:105029
94. Ibtehaz N et al (2022) PPG2ABP: Translating photoplethysmogram (PPG) signals to arterial blood pressure (ABP) waveforms. *Bioengineering* 9(11):692
95. Harfiya LN, Chang C-C, Li Y-H (2021) Continuous blood pressure estimation using exclusively photoplethysmography by LSTM-based signal-to-signal translation. *Sensors* 21(9):2952
96. Brophy E et al (2021) Estimation of continuous blood pressure from ppg via a federated learning approach. *Sensors* 21(18):6311
97. Yosinski J et al (2014) How transferable are features in deep neural networks? *Adv Neural Inform Process Syst* 27
98. Morassi Sasso A et al (2020) HYPE: Predicting blood pressure from photoplethysmograms in a hypertensive population. In: *International conference on artificial intelligence in medicine*, Springer
99. Wang W et al (2021) Cuff-less blood pressure estimation from photoplethysmography via visibility graph and transfer learning. *IEEE J Biomed Health Inform* 26(5):2075–2085
100. Mou H, Yu J (2022) Transfer learning with DWT based clustering for blood pressure estimation of multiple patients. *J Comput Sci* 64:101865
101. Leitner J, Chiang P-H, Dey S (2021) Personalized blood pressure estimation using photoplethysmography: A transfer learning approach. *IEEE J Biomed Health Inform* 26(1):218–228
102. Zhang Y et al (2022) A Refined Blood Pressure Estimation Model Based on Single Channel Photoplethysmography. *IEEE J Biomed Health Inform* 26(12):5907–5917
103. Ali NF, Atef M (2022) LSTM Multi-Stage Transfer Learning for Blood Pressure Estimation Using Photoplethysmography. *Electronics* 11(22):3749
104. Huthart S et al (2020) Advancing PPG signal quality and know-how through knowledge translation-from experts to student and researcher. *Front Digital Health*
105. Athaya T, Choi S (2020) Evaluation of Different Machine Learning Models for Photoplethysmogram Signal Artifact Detection. In: *2020 International conference on information and communication technology convergence (ICTC)*, IEEE
106. Zhang O et al (2021) Explainability Metrics of Deep Convolutional Networks for Photoplethysmography Quality Assessment. *IEEE Access* 9:29736–29745
107. Naeini EK et al (2019) A real-time PPG quality assessment approach for healthcare Internet-of-Things. *Procedia Comput Sci* 151:551–558
108. Roh D, Shin H (2021) Recurrence Plot and Machine Learning for Signal Quality Assessment of Photoplethysmogram in Mobile Environment. *Sensors* 21(6):2188
109. Gambarotta N et al (2016) A review of methods for the signal quality assessment to improve reliability of heart rate and blood pressures derived parameters. *Med Biological Eng Comput* 54(7):1025–1035
110. Desquins T et al (2022) A Survey of Photoplethysmography and Imaging Photoplethysmography Quality Assessment Methods. *Appl Sci* 12(19):9582
111. Lokendra B, Puneet G (2022) AND-rPPG: A novel denoising-rPPG network for improving remote heart rate estimation. *Comput Biol Med* 141:105146
112. Lu H, Han H, Zhou SK (2021) Dual-GAN: Joint BVP and Noise Modeling for Remote Physiological Measurement. In: *2021 IEEE/CVF Conference on computer vision and pattern recognition (CVPR)*, Nashville, TN, USA, 2021, pp 12399–12408. <https://doi.org/10.1109/CVPR46437.2021.01222>
113. Amelard R, Clausi DA, Wong A (2016) Spectral-spatial fusion model for robust blood pulse waveform extraction in photoplethysmographic imaging. *Biomed Optics Express* 7(12):4874–4885
114. Revanur A et al (2021) The first vision for vitals (v4v) challenge for non-contact video-based physiological estimation. In: *Proceedings of the IEEE/CVF international conference on computer vision*
115. Zahid Hasan SRR, Roy N (2021) MPSC-rPPG Dataset. <https://doi.org/10.21227/ddgz-tx88> IEEE Dataport
116. Kopeliovich M, Petrushan M (2022) rPPG Dataset. <https://doi.org/10.17605/OSF.IO/FDRBH>
117. Kachuee M, Kiani MM, Mohammadzade H, Shabany M (2015) Cuff-less high-accuracy calibration-free blood pressure estimation using pulse transit time. In: *2015 IEEE International symposium on circuits and systems (ISCAS)*, Lisbon, Portugal, pp 1006–1009. <https://doi.org/10.1109/ISCAS.2015.7168806>
118. Johnson A, Bulgarelli L, Pollard T, Horng S, Celi LA, Mark R (2022) MIMIC-IV (version 2.1). *PhysioNet*. <https://doi.org/10.13026/rrgf-xw32.2022>
119. Lee H-C et al (2022) VitalDB, a high-fidelity multi-parameter vital signs database in surgical patients. *Scientific Data* 9(1):279
120. Rogers P, Wang D, Lu Z (2021) Medical information mart for intensive care: a foundation for the fusion of artificial intelligence and real-world data. *Front Artif Intell* 4:691626
121. Data MC, Mark R (2016) The story of MIMIC. *Secondary Anal Electronic Health Records* 43–49

122. Zhang K et al (2016) Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Process Lett* 23(10):1499–1503
123. Jiang H, Learned-Miller E (2017) “Face Detection with the Faster R-CNN,” 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017), Washington, DC, USA, pp 650–657. <https://doi.org/10.1109/FG.2017.82>
124. Ye B, Shi Y, Li H, Li L, Tong S (2021) “Face SSD: A Real-time Face Detector based on SSD,” 2021 40th Chinese Control Conference (CCC), Shanghai, China, pp 8445–8450. <https://doi.org/10.23919/CCC52363.2021.9550294>
125. Garg D et al (2018) A deep learning approach for face detection using YOLO. In: 2018 IEEE Punecon, IEEE
126. Henriques JF et al (2014) High-speed tracking with kernelized correlation filters. *IEEE Trans Pattern Anal Machine Intell* 37(3):583–596
127. Badrinarayanan V, Kendall A, Cipolla R (2017) Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans Pattern Anal Machine Intell* 39(12):2481–2495
128. Simonyan K, Zisserman A (2015) Very Deep Convolutional Networks for Large-Scale Image Recognition, 3rd International Conference on Learning Representations (ICLR 2015), Computational and Biological Learning Society pp 1–14
129. He K, Zhang X, Ren S, Sun J (2016) Deep Residual Learning for Image Recognition, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, pp 770–778. <https://doi.org/10.1109/CVPR.2016.90>
130. Topiwala A, Al-Zogbi L, Fleiter T, Krieger A (2019) “Adaptation and Evaluation of Deep Learning Techniques for Skin Segmentation on Novel Abdominal Dataset”. In: 2019 IEEE 19th International conference on bioinformatics and bioengineering (BIBE), Athens, Greece, pp 752–759. <https://doi.org/10.1109/BIBE.2019.00141>
131. Krizhevsky A, Sutskever I, Hinton GE (2017) Imagenet classification with deep convolutional neural networks. *Commun ACM* 60(6):84–90
132. Gesche H et al (2012) Continuous blood pressure measurement by using the pulse transit time: comparison to a cuff-based method. *European J Appl Physiol* 112(1):309–315
133. Zhao Y, Huang Z, Gong L, Zhu Y, Yu Q, Gao Y (2023) Evaluating the impact of data transformation techniques on the performance and interpretability of software defect prediction models. *IET Software* 2023:6293074. <https://doi.org/10.1049/2023/6293074>
134. Bland JM, Altman DG (1995) Comparing methods of measurement: why plotting difference against standard method is misleading. *The Lancet* 346(8982):1085–1087
135. Yaacob H, Hossain F, Shari S, Khare SK, Ooi CP, Acharya UR (2023) Application of artificial intelligence techniques for brain-computer interface in mental fatigue detection: A systematic review (2011–2022). *IEEE Access* 11:74736–74758. <https://doi.org/10.1109/ACCESS.2023.3296382>
136. Khare SK, March S, Barua PD, Gadre VM, Acharya UR (2023) Application of data fusion for automated detection of children with developmental and mental disorders: A systematic review of the last decade. *Inform Fusion* 99. <https://doi.org/10.1016/j.inffus.2023.101898>
137. Khare SK, Acharya UR (2023) Adazd-Net: Automated adaptive and explainable Alzheimer’s disease detection system using EEG signals. *Knowl-Based Syst* 278. <https://doi.org/10.1016/j.knosys.2023.110858>
138. Khare SK, Acharya UR (2023) An explainable and interpretable model for attention deficit hyperactivity disorder in children using EEG signals. *Comput Biol Med* 155. <https://doi.org/10.1016/j.combiomed.2023.106676>

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Authors and Affiliations

Osama A. Omer<sup>1</sup> · Mostafa Salah<sup>2</sup>  · Loay Hassan<sup>1</sup> · Ahmed Abdelreheem<sup>3</sup> · Ammar M. Hassan<sup>4</sup>

Osama A. Omer  
omer.osama@aswu.edu.eg

Loay Hassan  
llhhaf@yahoo.com

Ahmed Abdelreheem  
a.abdelreheem@eng.svu.edu.eg

Ammar M. Hassan  
ammar@aast.edu

<sup>1</sup> Faculty of Engineering, Aswan University, Aswan, Egypt

<sup>2</sup> Faculty of Engineering, Sohag University, Sohag, Egypt

<sup>3</sup> Electrical Engineering Dept., South Valley University, Qena, Egypt

<sup>4</sup> Computer Science Dept., Arab Academy for Science, Technology and Maritime Transport, South Valley Branch, Aswan, Egypt