



A CNN based multifaceted signal processing framework for heart rate proctoring using Millimeter wave radar ballistocardiography

Rafid Umayer Murshed^{a,b,*}, Md. Abrar Istiak^a, Md. Toufique Rahman^a, Zulqarnain Bin Ashraf^a, Md. Saheed Ullah^a, Mohammad Saquib^b

^a Department of Electrical and Electronic Engineering (EEE), Bangladesh University of Engineering and Technology (BUET), Dhaka 1205, Bangladesh

^b Department of Electrical and Computer Engineering, The University of Texas at Dallas (UT Dallas), Richardson, TX, 75080, USA

ARTICLE INFO

Keywords:

Cardiovascular disease
Contactless measurement
Heart rate
Inter-beat interval
mm-wave radar
Convolutional neural network

ABSTRACT

The recent pandemic has refocused the medical world's attention on the diagnostic techniques associated with cardiovascular disease. Heart rate provides a real-time snapshot of cardiovascular health. A more precise heart rate reading enables a better understanding of cardiac muscle activity. Although many existing diagnostic techniques are approaching the limits of perfection, there remains potential for further development. In this paper, we propose MIBINET, a novel multifaceted approach for real-time proctoring of heart rate from Millimeter wave (mm-wave) radar ballistocardiography signals via inter-beat-interval (IBI) using a convolutional neural network (CNN). The central theme of our approach is to synergize the feature extraction capabilities of CNN with novel signal processing techniques, resulting in enhanced estimation accuracy while simultaneously reducing computational complexity. This proposed network can be used in hospitals, homes, and passenger vehicles due to its lightweight and contactless properties. It employs classical signal processing prior to fitting the data into the network. Although MIBINET is primarily designed to work on mm-wave signals, it is found equally effective on signals of various modalities such as PCG, ECG, and PPG. Our approach outperforms state-of-the-art techniques by more than 5% in inter-beat-interval (IBI) estimation accuracy. The architecture achieves a 98.73% correlation coefficient and a 20.69 ms Root-Mean-Square Error (RMSE) over 11 different test subjects. The paper contributes by being the first to apply CNN-based feature extraction in concert with unique signal processing strategies to mm-wave radar data for heart rate monitoring. Our methodology also introduces a synthetic IBI augmentation technique, custom loss function, and novel post-processing methods, all contributing to the robust performance of the model in various settings and modalities.

1. Introduction

Human vital signs like Heart Rate (HR), Heart Rate Variability (HRV), Respiration Rate (RR), and Oxygen Saturation (SpO₂) are important physiological indicators that reflect the physical and mental well-being of the human body. The heart pumps oxygenated and nutrient-rich blood all over the body. As the cardiac output is intimately associated with HR and stroke volume, HR is central to the cardiovascular process. HR measurement is crucial for a health monitoring system [1]. According to WHO Global Health Estimates, heart diseases such as myocardial infarction (MI), sudden cardiac death, heart attack, coronary artery disease, arrhythmia, heart valve disease, and heart infection have been the leading causes of death over the last few decades. In 2020, about 697,000 people died alone from heart disease in the United States; one in every five deaths [2,3]. Clinical biofeedback practice heavily emphasizes the control of cardiac

dynamics. To analyze biofeedback, HR, the number of heartbeats per minute is the most often measured metric. However, a healthy heart does not beat uniformly; it changes its rhythm with each beat. While HR focuses on the average number of beats per minute, HRV measures the specific variations in time between successive heartbeats. The period between heartbeats is measured in milliseconds (ms) and is referred to as the *R-R interval* or the *inter-beat interval (IBI)*. The sympathetic and parasympathetic branches of the autonomic nervous system (ANS) control HRV. Especially for diabetic and post-infarction patients, it is a crucial parameter in analyzing the behavior of the sympathetic and parasympathetic functions of the ANS [4]. An effective way of estimating HRV is measuring the variation in IBI [5], which can be used to detect probable cardiovascular disorders [6,7].

Electrocardiogram (ECG) [8,9] and Photoplethysmography (PPG) are among the recognized methods for measuring HRV indices [10].

* Corresponding author at: Department of Electrical and Computer Engineering, The University of Texas at Dallas (UT Dallas), Richardson, TX, 75080, USA.
E-mail address: rafidumayer.murshed@utdallas.edu (R.U. Murshed).

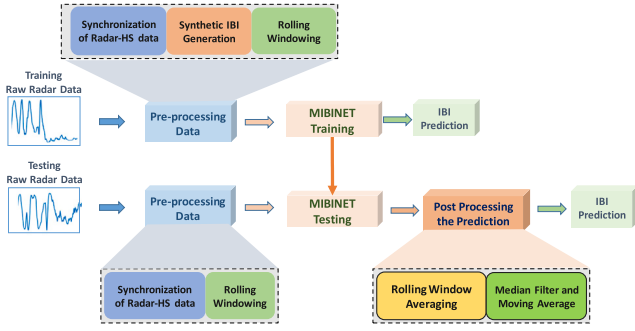


Fig. 1. Overall workflow of the proposed approach.

These systems require the patients to be in constant communication with humans. However, this can be challenging for remote patient monitoring, especially for elderly people living alone or recently discharged patients. Another disadvantage of the aforementioned systems is the limitation of mobility and freedom of the patients. In addition, the electrodes used by ECG and PPG may not only cause unnecessary discomfort but also potentially exacerbate symptoms. Moreover, the palpation of the probes could lead to false alarms and cause alarm fatigue [11,12]. Radar-based contactless tele-measurements offer a very comfortable means of continuous vital signal monitoring, which is essential for detecting early onsets of cardiovascular diseases. In this paper, we pursue that line of inquiry.

Millimeter wave (mm-wave) based technologies are ideal alternatives for contactless, continuous measurement of human vital signals [13]. The shorter wavelengths of mm-waves (30–300 GHz) allow for greater spatial resolution. To this end, a number of approaches have been explored to implement mm-wave systems in different biological and medical applications. The most notable applications include contactless measurement of arterial pulses [14], cancer diagnosis [15,16], and dental diagnosis [17]. In addition, the use of high-resolution mm-wave array beamformers has increased in medical imaging, gesture recognition, and navigation in recent years. Non-thermal, low-intensity electromagnetic radiation is used in mm-wave therapy (MWT), a novel and revolutionary method [18] of treating patients. The implementation of mm-wave FMCW radar [19] is used to measure breathing and HRs. To enhance the effectiveness and feasibility of these technologies in real-time, their implementation must take into account both computational efficiency and optimal precision. The above studies do not analyze both the computational efficiency and optimal precision of the used mm-wave technologies simultaneously. Unlike those, our proposed mm-wave-based technique, MIBINET, can estimate the IBI values with optimal precision in real-time.

To the best of our knowledge, MIBINET is the first neural network-based approach that operates on pre-processed mm-wave data to estimate the instantaneous IBI. Our work's fundamental premise is to improve estimation accuracy and computational complexity by fusing the feature extraction capabilities of CNN with distinctive signal processing strategies. It adopts a rolling window-based pre-processing approach to confine the input data to a predetermined size for conveniently feeding into a neural network's input layer. The network's output is also reprocessed with a novel rolling window averaging approach followed by various traditional post-processing filters to improve the estimation accuracy; see Fig. 1. Fig. 1 demonstrates the whole workflow from the raw radar signal to IBI prediction. Our numerical results demonstrate that the proposed MIBINET is capable of outperforming the state-of-the-art techniques in terms of IBI estimation precision by more than 5%. Below is a summary of the other major contributions of this investigation:

1. We introduce a synthetic IBI augmentation technique to enrich the dataset, significantly enhancing the correlation coefficient and reducing the root-mean-square error (RMSE). This

data augmentation enables our method to demonstrate robust performance across various patients and even in anomalous cases.

2. Our approach combines a rolling window-averaging technique with various traditional post-processing filters to improve the estimation accuracy, resulting in a more than 5% increase in IBI estimation precision compared to state-of-the-art techniques.
3. We employ a custom loss function specifically designed to reduce outliers, which leads to substantial improvements in model performance. Our designed lightweight 1D-CNN model architecture facilitates real-time use and exhibits high performance across 11 different test subjects.

The rest of this paper is organized as follows: Section 2 contains a review of the current state of contact-free vital sign monitoring. MIBINET is proposed in Section 3, along with the description of pre and post-processing techniques. Section 4 includes the numerical results. Finally, a brief discussion of the findings, limitations, and strengths of MIBINET is provided in Section 5. Regarding notation, scalars and vectors are represented in lower and bold lower cases, respectively. The n th element of vector \mathbf{a} is denoted by $\mathbf{a}(n)$.

2. Related works

A significant amount of research has focused on non-invasive contactless monitoring of vital signs (e.g., HR, RR, HRV) due to its many advantages. Non-invasive and contactless optical approaches for heart-beat monitoring based on optical Doppler interferometry and laser have been proposed in [20,21]. Here, the photo-EMF pulsed laser vibrometer (PPLV) is studied, where the subjects were instructed to first exhale and then hold their breath for as long as they could after inhaling. This is to avoid any potential muffling of the cardiac signals by the subjects' respirations. However, it might cause unnecessary discomfort to the patients. Fengyu Wang et al. [22] proposed contactless HRV monitoring using mm-wave radio. First, they developed a user-locating target detector without calibration. Heartbeat signal extractors then optimize the decomposition of chest-movement-modulated channel information to uncover the desired signal. Now the pulse signal's peak position can be used to evaluate HRV parameters for each target utilizing the IBI values. This method, mmHRV, can assess HRV with a median IBI error of 28 ms (w.r.t 96.16% accuracy) for 11 players in the line of sight (LOS). For non-LOS, it is 31.71 ms. However, this proposed method exhibited considerably uneven errors among the participant subjects. Zhang et al. [23] suggested radio signal-based contactless MI detection. This work establishes MI detection using RF signals, providing contactless, non-intrusive, continuous home monitoring for MI hazards. They also proposed heartbeat signal segmentation and MI detection algorithms. Extensive evaluations have been conducted to confirm the effectiveness of Health-Radio. However, this approach has not been generalized to work over different signal modalities such as PPG, ECG, and PPG.

Radar technology is one of the most promising possibilities for contactless, non-invasive monitoring of vital signs such as cardiac signals. Impulse radio ultra-wideband (IR UWB) radar was used to monitor vital signs [24,25]. While monitoring, this IR UWB-based technique exploits signal properties, completely disregarding any object qualities with which this signal interacts. Body-coupled antennas and UWB pulsed radar in-body monitoring of lungs and heart motion are used in this scenario. The radar has to be installed at a specific location. Moreover, sudden oscillations cause abrupt phase fluctuations, which impact HR predictions. Many methods for calculating HR [26,27] from UWB radar data have recently been developed. Their performance, however, is still insufficient for practical applications. Frequency-modulated continuous wave (FMCW) radar [24], continuous wave (CW) Doppler radar for HR monitoring, and respiration monitoring have been developed significantly during the past few years [28–30]. Because CW radar or Doppler

radar does not capture the target's range or distance, FMCW radar is designed to overcome this issue due to its range and radial velocity measurement. Since FMCW radar operates at lower transmit power, the received signal will not only be distorted by the environment but also could be weak. Although beamforming and range-gating approaches can separate the signal of interest from the noise, several difficulties, such as random body motions, must be addressed before radar-based non-contact measurements can be implemented in real-world applications. Sakamoto, Takuya, et al. [31] addressed the use of a UWB radar system for estimating the human HR precisely. The performance of the proposed approach was demonstrated through measurements. Target classification and type recognition are achievable with UWB radar because the received signal contains information not just regarding the target as a whole but also about its individual component elements [32]. The suggested approach estimates the HR efficiently and correctly by using the feature points of a radar signal. Nonetheless, as heartbeat waveforms vary, even within the beat-to-beat interval, Fourier and periodicity-based approaches are ineffective for estimating instantaneous HR in real time. Consequently, a Wavelet-Transform (WT) based method [33] has been proposed for faster HR detection using Doppler radar because of the insufficient frequency resolution of the Fourier transform (FT). Unfortunately, the WT approach requires extensive signal processing in order to identify HR.

In 2010, researchers started applying machine learning (ML) in the field of ballistocardiography signals. Bruser et al. [34] brought the idea of an unsupervised modified K-means clustering training algorithm to estimate parameters for the BCG signal. Then, the parameters followed by the heartbeat estimation and refinement process combine three indicators for localizing the heartbeat. Collectively, the method called BEAT has to be re-trained when facing a significant change in the BCG signal. Later, a real-time approach for identifying individual heartbeats without requiring intensive signal processing was studied in [35]. It is an Artificial Neural Network (ANN) based Heartbeat Detection Technique using a state-of-the-art mm-wave radar sensor. The ANN is trained on the raw radar signal, offering computational simplicity while estimating IPIs (inter-pulse intervals) with relatively low accuracy. Hence, a signal of high HRV is not suitable for this study. In addition, this shallow ANN model has a 2% chance of missing a pulse. Nowadays, adopting deep learning methods enables the proliferation of automatic, portable, non-invasive data-driven HR monitoring. In 2019, Dwaipayan et al. [36] proposed a temporal model CorNET (CNN + LSTM) method to predict HR and biometric identification. This was evaluated on two subjects and consequently lacked subject bias. This model needs longer training in order to adapt to the HR variability in an ambulant environment.¹ Xiangmao et al. [37] proposed a method to estimate HR after acquiring clean input PPG data from a denoising CNN (DCNN). Although DCNN adds more robustness to real-life artifacts, it adds overhead in time complexity. On a different note, the first triumphant attempt to predict pulse rate (PR) from facial video data after the surge of computer vision was the proposition of PRnet [38]. This method leverages the synergy of 3D convolution and LSTM to predict pulse rate from spatiotemporal features with less error. It requires a minimum of 2 s (60 frames) to effectively measure PR, which limits the method from being a real-time pragmatic solution. Subsequently, a method was developed for identifying heart problems in an embedded system using an energy-efficient CNN [39]. This method involved aggregating retrieved characteristics from segmented parts of the electrocardiogram (ECG) data. Despite its great performance and energy efficiency, this network does not possess the capability to capture local characteristics of ECG signals.

In addition, a cutting-edge motion artifact removal method for Imaging Ballistocardiography (iBCG), providing a non-contact heart

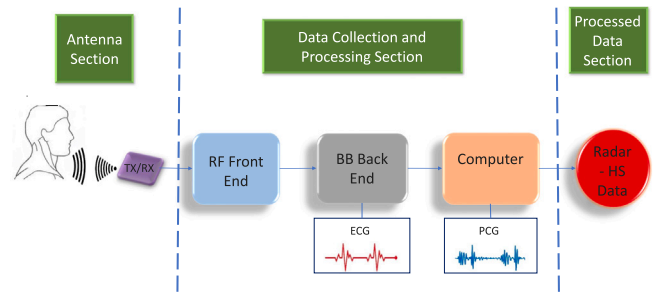


Fig. 2. Data collection scheme — The figure depicts how the data was collected.

rate measurement approach, has recently emerged [40]. This technique involves reconstructing Z-axis signals and applying adaptive filtering, PCA, and CCA techniques to mitigate rigid and non-rigid motion artifacts effectively. While rigorous experiments confirm its exceptional performance, especially in the presence of motion artifacts, challenges remain, including potential sensitivity to substantial motion and the need for further research on automating threshold settings.

Further, The Radar-Beat system offers a holistic approach to non-invasive heartbeat monitoring through mmWave FMCW radar technology [41]. It leverages radio-frequency tech to detect body surface micro-vibrations and extract heartbeats, incorporating sensitive algorithms for motion detection, optimized range-bin selection, personalized templates, and a global optimization model for accurate heartbeat duration estimation. While it shows strong agreement with synchronized ECG devices, especially during extended monitoring in diverse positions, it lacks a thorough examination of the impact of real-world noise. Further research is needed to assess its robustness in varying conditions and with specific health conditions. Additionally, the Radar-Beat system reliance on Gaussian distributions and loss weights in its optimization model may not universally apply, necessitating further validation.

Unlike previous studies, which often rely on static, pre-defined datasets, our approach allows for greater adaptability and versatility. Our synthetic IBI Generation Augmentation technique, in particular, proves more effective than existing methods for handling irregularities in heart rate and other cardiac conditions. This novelty is especially critical in real-world applications, where heart rate conditions are often far from ideal.

3. Methodology

In this section, we describe the working principle of MIBINET. Since it is a DNN-based technique, it needs to be trained and validated on labeled data; we describe how it is collected.

3.1. Data acquisition

Fig. 2 is a representation of the data collecting and processing procedure following the work presented in [42]. Shi et al. [42] made the dataset available in 2020 to foster the field of radar heart sound (Radar-HS) in a contactless manner. The dataset was approved by the ethics committee of the Friedrich-Alexander-Universität Erlangen-Nürnberg for maintaining guidelines and regulations. Their tailored hardware setup initiates from the RF front end, as shown in Fig. 2, consisting of a six-port, and it is utilized as a quadrature interferometer for radar applications. The six-port has two input signals and four output signals, where two input signals consist of a reference signal at a defined frequency and a received signal reflected from the target. The antenna direction was set perpendicular to the test subject's thorax surface to maximize the signal quality. The baseband board back end (BB Back End) receives radar signals from the RF front end and digitizes

¹ Unlike [36], we tackled variation with augmentation and rolling window averaging techniques.

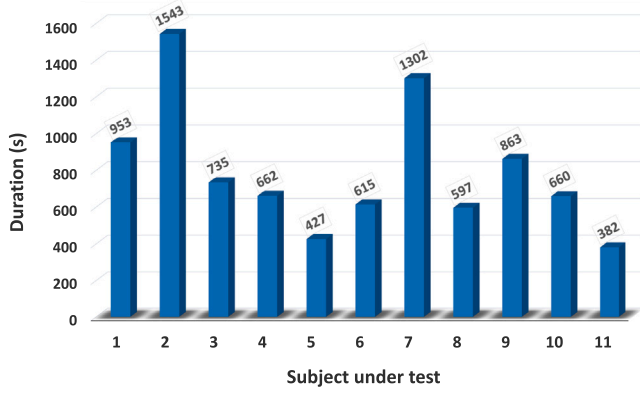


Fig. 3. Total signal duration of each subject in the dataset.

the signals. It is equipped with an ECG and respiration sensor (RS) for simultaneous sampling. The digitized raw signals were subsequently received by a computer equipped with a PCG that served as a reference sensor. PCG signals are stored after re-sampling and synchronization.

Let us now take a brief glance at the processed data, which consists of the raw radar data along with the filtered data from the respiration sensor, the ECG leads, the PCG data, and the Radar heart sounds. Fig. 3 provides a breakdown of the collected data. As we can see, 11 subjects participated in this measurement process. Although the total data consists of 13376 s of recordings, only the default scenarios are of interest to us as Shi et al. [42] made them publicly available. In the following sub-section, we describe our adopted rolling window-based pre-processing scheme, which allows us to meta-morph the collected dataset seamlessly in a form compatible with deep neural networks.

3.2. Dataset distribution

Following standard practice, the entire dataset was permuted into 11 folds, with each subject's data appearing only once in the test set of each fold. Each fold consists of separate training, validation, and test sets. The folds are numbered according to the ID of the subject in the test set. For instance, fold one indicates that the data of the first subject is in the test set. The training and validation sets consist of the data of the remaining ten users, with eight in training and two in validation.

The distribution of the IBI values from the data of different test subjects reveals the typical variations usually found in clinical trials. For example, as shown in Fig. 4a, the median values of the participant test subjects vary from 0.6 to 0.8 s (s), with subjects 3 and 10 having comparatively lower IBI values while subjects 7, 8, and 9 exhibit relatively high values.

Analyzing the distribution of each fold shows some disparities between the range of IBI values in the training and test sets. In particular, folds 3 and 10 display significant variations, making it quite challenging to train deep learning-based models. This leads us to propose a novel data augmentation approach to training the deep models, which will be elaborated on next.

3.3. Synthetic IBI generation augmentation

Due to the scarcity of lower IBI values in the training sets of folds 3 and 10 (from Fig. 4a), the model struggles to perform well on the test sets, which contain a relatively lower range of IBI values. We resolve this issue by generating synthetic radar-HS signals to inject smaller IBI values into the training set. To synthesize radar-HS signals characterized by smaller IBI values, specifically in the range of 0.5–0.6 s, we first start with signals that have IBI values in the higher range, preferably in the region of 0.9–1.1 s. We then left-shift those signals in the time domain by 0.4–0.5 s and add them with their shifted

versions as depicted in Fig. 5(a). Mathematically, it can be expressed as follows.

$$x_{\text{aug}}(t) = x_o(t) + x_o(t + \alpha), \quad (1)$$

where $x_{\text{aug}}(t)$ and $x_o(t)$ denote the augmented and original signals respectively. Here, α is a uniform random variable between 450 ms and 550 ms. Now the R-peaks of the augmented signal are the union of the R-peaks from the original signal and the left-shifted signal, as shown in Fig. 5(b). Due to such augmentation, the distribution of the resulting training set becomes much more consistent, as can clearly be observed in Fig. 4b. In brief, this IBI augmentation technique synthesizes comparatively smaller IBI values corresponding to elevated tachycardia heart rates. Similarly, right-shifting the signal can make the distribution adaptable for high IBI bradycardia conditions. It is important to note that this IBI augmentation technique is particularly effective for periodic signals, ensuring the correct deployment of the method while maintaining the accuracy of the results.

In order to further facilitate our neural network-based approach, we supplement the processed signals with an additional rolling window-based pre-processing scheme, which will be discussed next.

3.4. Rolling window-based pre-processing and augmentation

IBI is the time interval between two consecutive R-peaks, as shown in Fig. 6a. It can be noticed that the rolling window extends over 8 consecutive R-peaks of the processed signals, each containing 7 IBIs; see Fig. 6b. Although the specific size of the rolling window is not based on exact theoretical calculation, our empirical results demonstrate that the proposed size, $L = 7$, is near optimal for this study to trade-off between computational time and performance.

Typically, most neural networks accept predefined dimensions of the inputs and outputs. However, the available radar data does not adhere to those constraints. Hence, we restrain the dimensions of the inputs and outputs to a predefined value. Getting inspired from [43,44], by using a rolling window of size 8, we are able to restrict the size of the output to a vector of size 7. In our dataset, each input consists of a windowed version of the original data containing 7 IBIs, and the maximum windowed input length is found to be 4885 samples. In order to restrict the dimension of the input data, we zero-pad all the inputs to a size of 4910 samples. These zeroes are inserted randomly at the start and end of each window to remove any prospective bias during training. The starting point of every successive rolling window is randomly selected between two consecutive R-peaks to remove any regional bias.

3.5. Architecture of the MIBINET

Our proposed CNN architecture aims to balance computational efficiency and high prediction accuracy for estimating IBI values. The architecture, named MIBINET, incorporates fewer parameters than other contemporary networks while maintaining high accuracy. The rationale behind developing such an architecture is to create a lightweight and efficient real-time heart rate monitoring solution.

MIBINET combines standard spatial convolutional layers and depth-wise separable convolutional layers to extract distinguishable features from the one-dimensional input data. The intuition behind using convolutional layers first is to exploit their ability to capture local patterns and spatial dependencies in the input data, enabling the extraction of meaningful features. To ensure faster convergence during training, the input data is first normalized and passed through an initial batch normalization layer that adjusts the mean and variance of the input. Following this preprocessing step, the network is comprised of a sequence of convolutional and pooling layers with varying kernel numbers and sizes. We have determined these kernel parameters

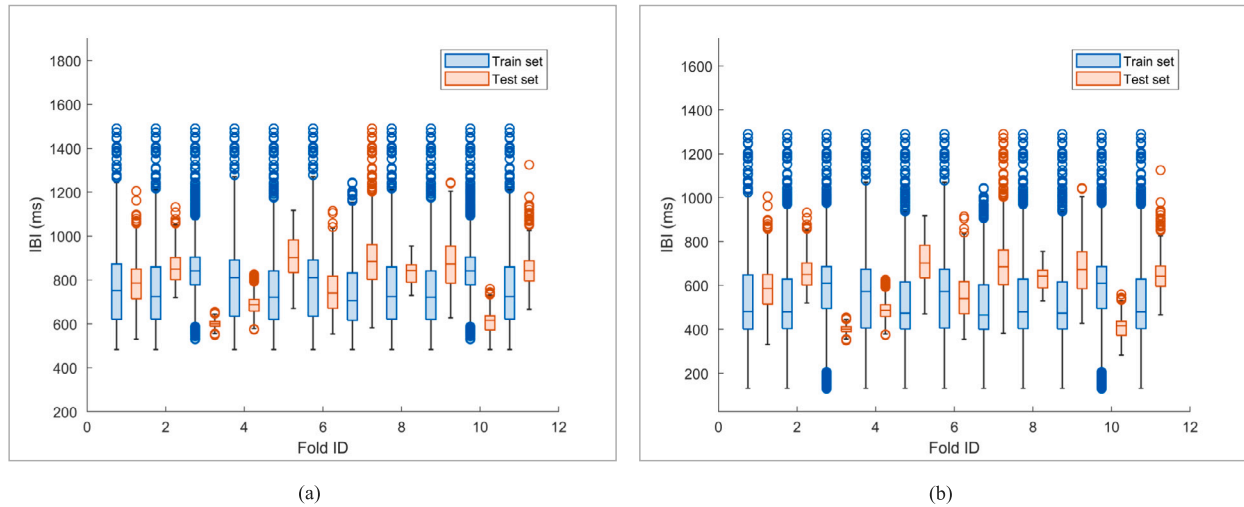


Fig. 4. Distribution of IBI values in box plot: (a) before augmentation and (b) after augmentation.

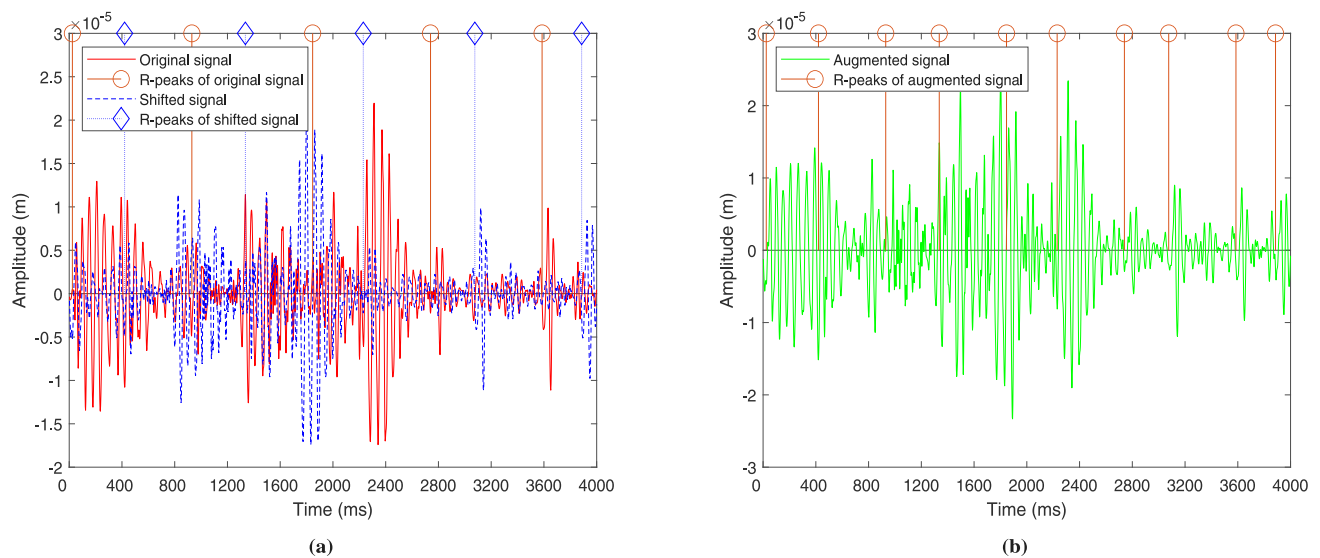


Fig. 5. Synthetic IBI augmentation: (a) superposition of signals for augmentation and (b) augmented signal with new R-peaks.

through extensive experimentation and trial-and-error to optimize feature extraction. The reasoning for employing pooling layers following the convolutional layers is to diminish the spatial dimensions of the feature maps while preserving crucial information, thereby enhancing computational efficiency and mitigating overfitting by reducing the number of parameters.

Subsequently, the extracted features are fed into a fully connected network of dense layers. The use of fully connected layers after the convolutional layers allows for the combination and processing of the extracted features at a higher level, ultimately leading to the final prediction. The final dense layer predicts an array of length seven, representing consecutive IBI values. We employ the swish activation function [45] in the convolutional layers, as it outperforms ReLU in this context, while ReLU is used in the fully connected dense layers. To further enhance computational efficiency, MIBINET utilizes multiple 1D depth-wise separable convolutional layers, reducing the number of parameters involved. A schematic diagram of the entire model architecture is illustrated in Fig. 7.

3.6. Custom loss function

The loss function plays a vital role in supervised learning algorithms such as feed-forward networks. It determines how the predictions approach the target labels. Since our goal is to measure the IBI values consistently with high precision, we not only require the error between the predictions and true values to be small but also want there to be a high correlation between the predicted values and the true values. To achieve this, we design a novel weighted loss function that takes into account the aforementioned considerations. The designed loss function is a weighted sum of three well-known loss functions for regression tasks, namely, the mean squared error (MSE) loss, the Huber loss (HL) [46], and the mean absolute error (MAE) loss. While the MSE is too sensitive to outliers, the MAE weights all the errors equally, disregarding outliers completely. The HL provides a good balance between both the MSE and the MAE. Finally, to ensure a high correlation between the predicted outputs and the actual value, we add another component to our weighted loss in the form of correlation coefficient loss. The resultant loss function can be expressed as

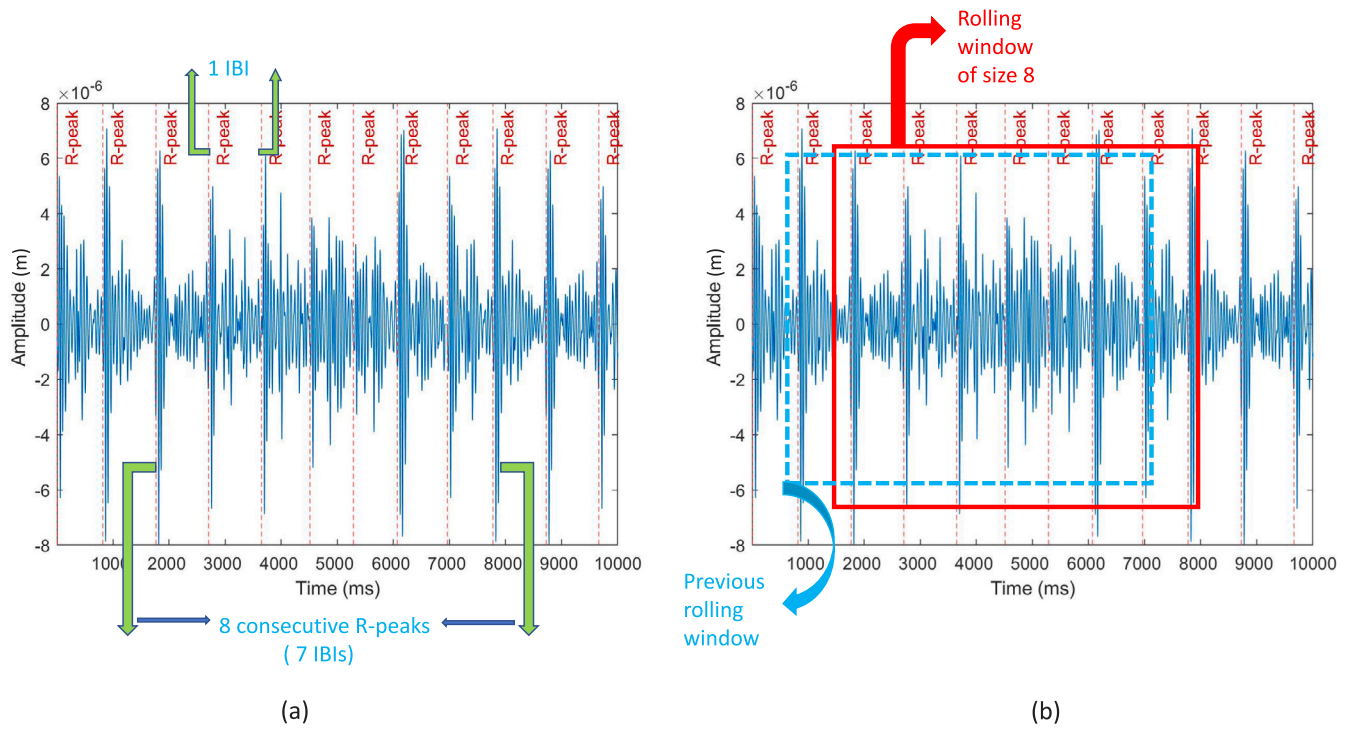


Fig. 6. Rolling window-based pre-processing: (a) original data and (b) windowing of data.

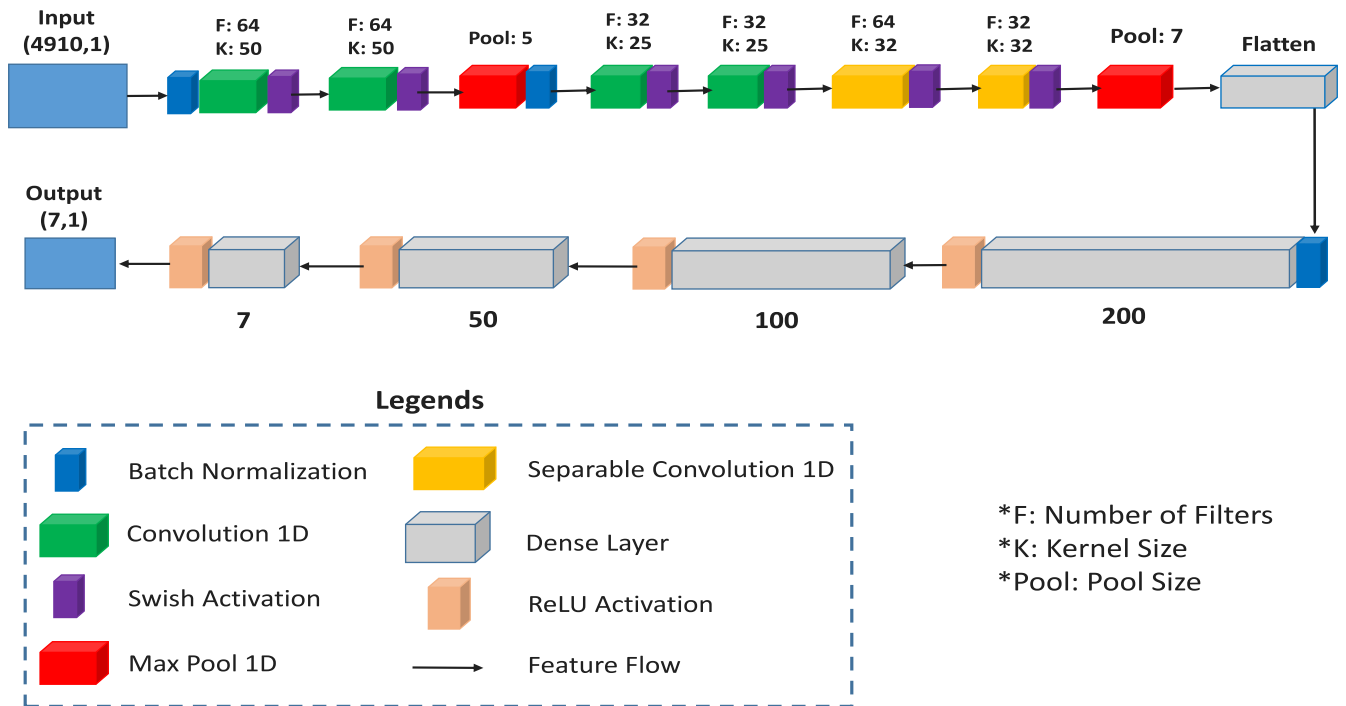


Fig. 7. Proposed MIBINETs' 1D-CNN architecture.

$$\mathcal{L}_{\text{corr}} = w_1 \times (1 - r^2) + w_2 \times L_{\text{HL}} + w_3 \times \epsilon^2 + w_4 \times |\epsilon| + w_5 \times L_{\text{HL}}^a \quad (2)$$

where L_{HL} denotes the well-known Huber-loss value, L_{HL} is a modified asymmetric version of the Huber-loss, r represents the Pearson correlation coefficient, ϵ is the error in the prediction, and w_1, w_2, w_3, w_4 and w_5 are the weights applied to the correlation loss, Huber-loss, MSE, MAE, and the asymmetric HL, respectively. The exact values used for

w_1, w_2, w_3, w_4 , and w_5 are determined through extensive trial and error to best suit the required task. It is also noteworthy that none of the well-known regression losses (i.e., MSE, MAE, mean absolute percentage error, etc.) except HL solely performed well in our experiments. Here, for each value ϵ in error = $y_{\text{pred}} - y_{\text{true}}$, L_{HL}^a is given by

$$L_{\text{HL}}^a = \begin{cases} 0.5 \times (\epsilon^2 + |\epsilon| + \rho \times \epsilon) & \text{if } |\epsilon| \leq \psi \\ 2 \times \psi \times |\epsilon| + \rho \times \psi \times \epsilon - 0.5 \times (\psi)^2 & \text{if } |\epsilon| > \psi, \end{cases} \quad (3)$$

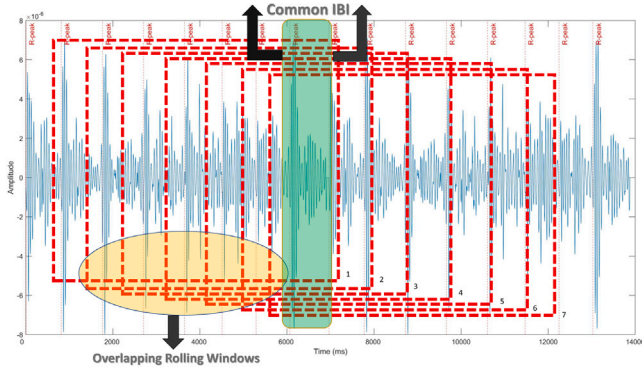


Fig. 8. Rolling window averaging: Illustration of seven overlapping successive windows and their intersection.

where ‘ ψ ’ is a constant parameter that determines the exact behavior of the loss towards outliers, and ρ is either -1 or $+1$, depending on whether we want positive or negative asymmetry. We choose a typical value of $\psi = 1$ for all simulations. The value of ρ is chosen to be $+1$ since the dataset had more high IBI values.

3.7. Post-processing

We first process the outputs by a rolling window averaging scheme to smooth out the predictions. This exploits our rolling window-based pre-processing scheme, where there is considerable overlapping between neighboring window samples. More precisely, each window retains seven R-peaks from the previous window and hence has exactly six of the same IBIs. This is utilized during the post-processing, where all predictions on the same initial IBI value are averaged, and hence the errors are further reduced. Following the post-processing schemes of [42], we also use a median filter of length 5 and a moving average filter of length 6, which boosts the reliability of the model’s predictions.

As described earlier, each rolling window consists of seven consecutive IBIs; see Fig. 6. Consequently, this leads to the overlapping of successive windows, and the intersection of each of seven such successive windows contains a single common IBI as shown in Fig. 8. Furthermore, our model predicts the IBI values from each window separately. Thus, it gives seven distinct predictions for each IBI value. Averaging those IBI values yields a more accurate prediction. Mathematically, this can be expressed as

$$z(k) = \frac{\sum_{i=1}^7 y_{k-7+i}(8-i)}{7}; \quad k = 7, 8, \dots \quad (4)$$

where $z(k)$ denotes the k th IBI value of a sample and y_j contains seven consecutive IBI predictions from the j th rolling window. The comprehensive approach of our proposed methodology is succinctly outlined by the pseudocode provided in Algorithm 1.

4. Experimental results

In this section, we discuss the experimental data and analysis used to evaluate the performance of our proposed MIBINET. We assess the efficacy of our proposed system using the Pearson correlation coefficient, and root means square error (RMSE). These metrics are chosen due to their suitability for regression tasks. In addition, the established state-of-the-art in [42] is used as the baseline methodology.

4.1. Evaluation metrics

The expression of the Pearson correlation coefficient is

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (5)$$

Algorithm 1: Pseudocode of MIBINET pipeline.

Input: Synchronized Radar HS of a subject, x_g

Output: Estimated IBI y_o

- 1: //Training
- 2: Step 1: Construct fixed length vector by Rolling Windowing R_L , where L = window length (= 7 IBIs in this work)
- 3: for $i = 1, 2, \dots, k_{max}$ (number of windows in a sample) do
- 4: $x_r[i] \leftarrow R_L(x_g)$;
- 5: end for
- 6: Step 2: Synthetic IBI generation augmentation, A (equation (1))
- 7: $x_a \leftarrow A(x_r)$;
- 8: Step 3: Fitting to DNN model, M
- 9: $\hat{y} \leftarrow M(x_a)$
- 10: Save M based on the lowest proposed loss (equation (2));
- 11: //Testing
- 12: Step 1: Repeat step 1 of training on test subject Radar HS
- 13: Step 2: Repeat step 3 of training and get \hat{y}
- 14: Step 3: Post-processing on prediction \hat{y}
- 15: $y_r \leftarrow RA(\hat{y})$, where, $RA(\cdot)$ is rolling averaging with L window length (equation (4));
- 16: $y_o \leftarrow F(y_r)$, where, $F(\cdot)$ is median filter followed by moving averaging;

where x_i and y_i are the i th predicted IBI and the ground truth IBI, respectively. Their sample means are denoted by \bar{x} and \bar{y} . The term n denotes the total number of IBIs in the sample. The other metric, RMSE, can be defined as

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_i - y_i)^2}{n}}. \quad (6)$$

It is to be mentioned that all 11 users’ ground truth and predicted IBI are concatenated prior to the evaluation of the final results.

Another regression metric named Coefficient of determination (R^2) is common to complement the correlation coefficient and RMSE metrics. R^2 measures the goodness of fit. As the range IBI for one subject is quite large, it mandates checking fitting capability. R^2 is also vulnerable to outliers and can indicate the overfitting issue for many folds. The R^2 can be measured by the formula:

$$R^2 = 1 - \frac{\sum_{i=1}^n (x_i - y_i)^2}{\sum_{i=1}^n (x_i - \bar{y})^2} \quad (7)$$

4.2. Training setup

All simulations are conducted using Google Colaboratory, a Python development environment that runs in the browser using Google Cloud and provides free access to powerful graphical processing units (GPU). Our proposed MIBINET and peripherals are implemented in Python 3.7 utilizing TensorFlow 2.9 and a Tesla T4 GPU provided by Google Colaboratory. All codes are executed with the same setup to enable an accurate and fair comparison of different processes. We have trained our network by utilizing a batch size of 1024 for 200 epochs. The learning rate is tuned using an exponential decay scheduler as shown in Fig. 9. The initial learning rate is set to 0.007. It is decreased by half every 40 epochs to 0.0018 after 80 epochs. Then, it is set to 0.00007 and is decreased by 10 times every 40 epochs through the next 120 epochs. All experiments used Adam [47] optimizer along with momentum with a decay of 0.9. In addition, the best weights were saved based on the validation-weighted metric. The weighted metric can be expressed as

$$\mathcal{M}_{\text{weighted}} = \alpha_1 \times (1 - r^2) + \alpha_2 \times \epsilon^2 + \alpha_3 \times |\epsilon|, \quad (8)$$

where the weights α_1 , α_2 , and α_3 are chosen to be 10, 0.1, and 0.1, respectively, to ensure equal importance on correlation coefficient and

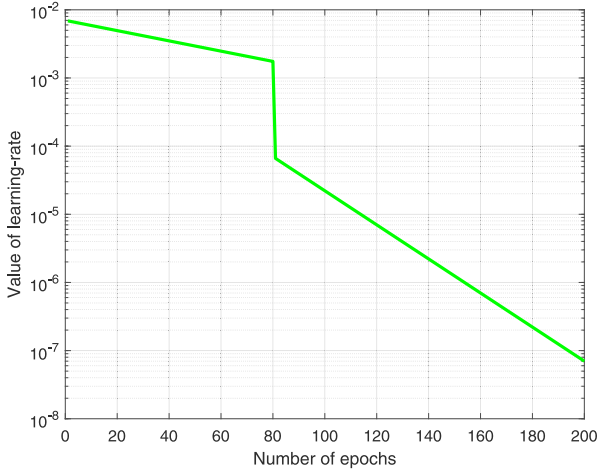


Fig. 9. Learning rate used for MIBINET training with the number of epochs.

Table 1

Simulation parameters.

Parameters	Values
Rolling window size	7
Input size	4910
Output size	7
Initial learning rate	0.007
Number of epochs	200
Batch size	1024
Pool size	5,7
Number of filters	32,64
Kernal size	25, 32, 50

Table 2

Fold-wise performance.

Test user no	Correlation coefficient, r (%)		RMSE (ms)	
	HSMM based [48]	MIBINET	HSMM based [48]	MIBINET
1	78.41	97.08	59.58	19.77
2	91.93	98.99	25.67	8.45
3	94.09	92.90	5.59	11.88
4	93.34	97.56	19.59	16.99
5	89.32	96.04	38.51	24.10
6	61.69	88.91	89.58	24.34
7	87.92	97.01	53.76	24.43
8	98.75	99.18	7.32	5.70
9	85.68	98.80	77.63	17.21
10	96.91	93.19	13.81	39.19
11	81.40	98.58	41.75	11.54

RMSE. Final model predictions are compared after the post-processing as mentioned in sub- Section 3.7. It is to be noted that, for every fold, the model was trained with the same parameters. In all our simulations, we have set $w_1 = 0.002$, $w_2 = 1.0$, $w_3 = 0.0096$, $w_4 = 0.002$, and $w_5 = 0.0032$ in (Eq. (2)) as those are found to be near optimal. Table 1 outlines several notable simulation parameters.

4.3. MIBINET vs. HSMM method

The Radar-HS dataset in [42] has recordings sampled at multiple frequencies. To maintain consistency, all the recordings are resampled to 500 Hz. We then prepare the dataset for MIBINET as described in Section 3.2. Following the aforementioned pre-processing and post-processing steps, the final results are generated, and the fold-wise results are given in Table 2. The fold-wise results of the HSMM method, the existing state-of-the-art, are also provided here for ease of comparison.

In Table 2, it can be noticed that apart from folds 3 and 10, our proposed MIBINET significantly outperforms the state-of-the-art HSMM

Table 3

Comparison between MIBINET and state-of-the-art HSMM algorithm.

Dataset	Methodology	r (%)	R^2	RMSE (ms)
Radar-HS [42]	HSMM	93.66	0.87	47.94
	MIBINET	98.73	0.97	20.69
PCG [42]	HSMM	87.02	0.76	53.52
	MIBINET	98.76	0.98	21.02
PTB-XL ECG [49]	HSMM	71.83	0.55	140.55
	MIBINET	99.60	0.98	11.93
Wrist PPG [51]	HSMM	-5.55	0.01	2641.43
	MIBINET	56.01	0.33	182.28

in both metrics. To substantiate the robustness and efficacy of our proposed approach compared to its counterpart, we test their performance over various datasets of different modalities. These particular ones are chosen because of their availability and authenticity. In the following subsections, we present a brief account of these datasets.

4.3.1. PCG dataset

Shi et al. [42] also provide the phonocardiograph (PCG) signals corresponding to the Radar-HS data from the previously considered 11 test subjects. With the similar split, recording number, sampling rate, and ground truth of Radar-HS, it just differs in the signal modality.

4.3.2. PTB-XL ECG dataset

A 12-lead electrocardiography [49] dataset comprising 21,837 records from 18,885 patients is used for comparison. The sampling rate is 100 Hz in the given recordings. However, we resample the dataset to 500 Hz to maintain conformity with the Radar-HS dataset. As it has a massive number of unique subjects under test, it can test the subject dependency of an algorithm. Hence, this dataset is used to demonstrate the universality of MIBINET. For dataset validity, it is split sequentially into train, validation, and test sets with a ratio of 60:20:20 based on the number of subjects. For the sake of simplicity, only the signal from lead II is considered, as the QRS complex here is more prominent compared to the ones in the other leads. It is to be noted that ECG signals are annotated using BioSPPy [50] as suggested in [49].

4.3.3. Wrist PPG

Jarchi et al. [51] released a PPG dataset collected using smart-watches from 8 participants. Here, all the recordings are sampled at a 256 Hz sampling rate, and 19 recordings are recorded in four different conditions, i.e., running, walking, fast-easy bike riding, and slow-difficult bike riding. We include this dataset as it has heavy motion artifacts, typical in real-life scenarios. Here, the R-peaks are determined from the given synchronized ECG signal. It also contains high HRV cases due to sudden initiation of motion and abrupt stoppages. Utilizing the 19 recordings, 19-fold cross-validation results are reported where fifteen, three, and one folds are in training, validation, and testing, respectively.

Now, we present the cumulative performance of MIBINET and HSMM in the above four (Radar-HS, PCG, ECG, and PPG) datasets; see Table 3. Here, we observe the following:

1. For the Radar-HS dataset, MIBINET provides an overall performance improvement of more than 5% in terms of correlation and 27 ms in RMSE, compared to HSMM.
2. In the PCG dataset, our algorithm performed even better than in Radar-HS. A higher correlation coefficient and lower RMSE portrays this performance improvement.
3. For the ECG dataset, the performance improvement of MIBINET over HSMM is even more prominent compared to the previous two datasets (correlation coefficient improvement > 27%).

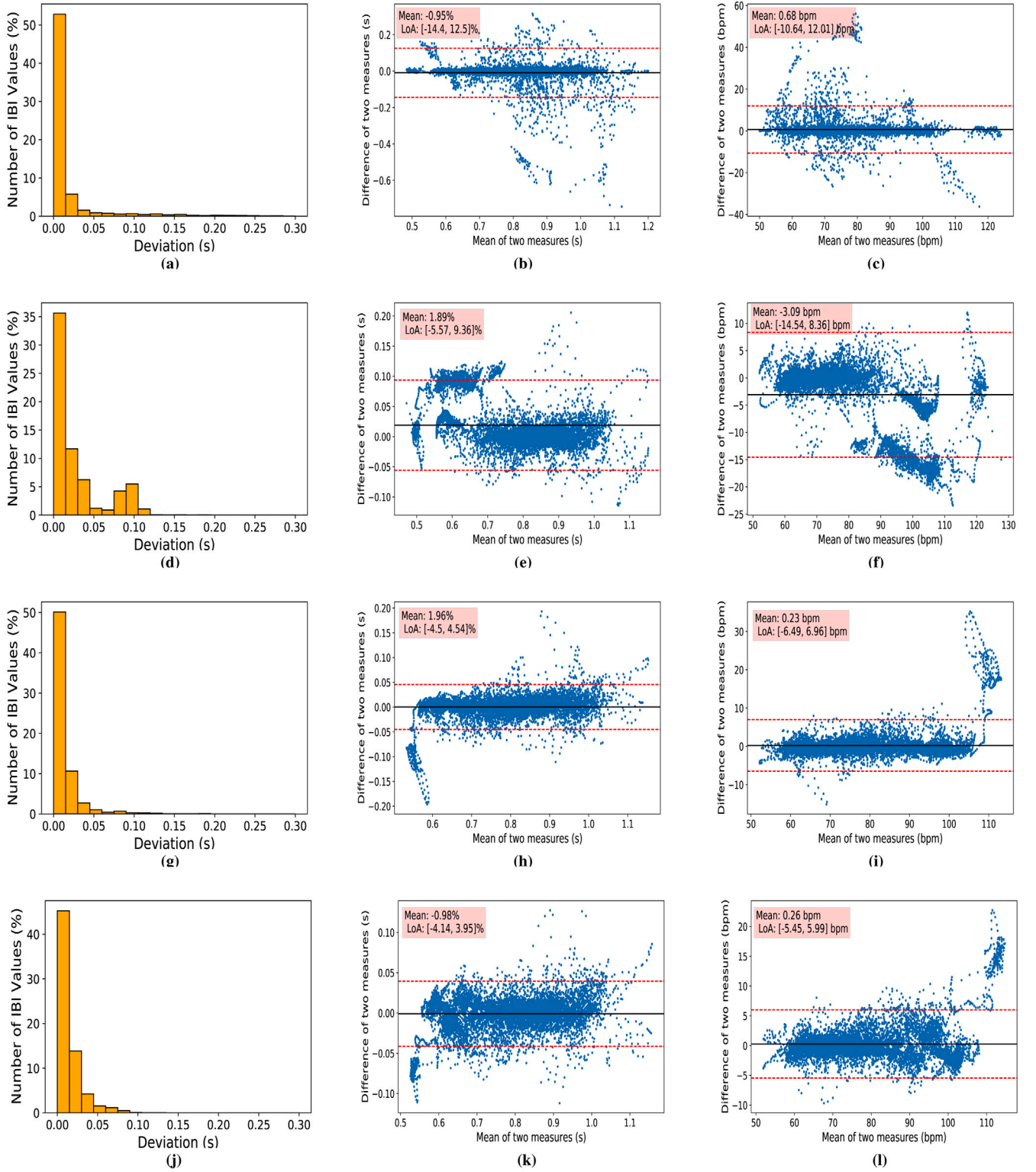


Fig. 10. Performance comparison based on error: (a)–(c) are plots for state-of-the-art HSMM-based algorithm. (d)–(f) are plots for Huber loss with augmentation. (g)–(i) are plots for proposed loss without augmentation. (j)–(l) are for proposed loss with augmentation. In the first column, error deviation plots; in the second column, Bland Altman plots of deviation between ground truth (Radar-HS) and predicted values; and in the third column, Bland Altman plots of deviation in the BPM unit are depicted.

4. Although the PPG signal is more challenging due to its miscellaneous artifacts and high noise content, MIBINET still outperforms the HSMM by a massive margin. The results in the final row of Table 3 show the robustness of our algorithm in challenging conditions induced by this PPG modality.

4.4. Performance of proposed loss function

A major contribution of this work is the weighted custom loss introduced in (Eq. (2)). The efficacy of the proposed loss is tested on various modalities of the dataset. We compare the performance of the proposed weighted loss with that of the best-performing existing

Table 4

Performance comparison of proposed loss.

Dataset	Loss function	r (%)	R^2	RMSE (ms)
Radar-HS [42]	Huber	97.31	0.94	42.55
	Proposed	98.73	0.97	20.69
PCG [42]	Huber	93.36	0.88	60.85
	Proposed	98.76	0.98	21.02
PTB-XL ECG [49]	Huber	37.90	0.14	248.57
	Proposed	99.60	0.99	11.93
Wrist PPG [51]	Huber	52.56	0.29	1117.47
	Proposed	56.01	0.33	182.28

Table 5

Performance comparison of augmentation on MIBINET.

Dataset	Augmentation	r (%)	R^2	RMSE (ms)
Radar-HS [42]	–	98.44	0.96	23.08
	✓	98.73	0.97	20.69
PCG [42]	–	88.46	0.78	61.03
	✓	98.77	0.98	21.02

regression loss, namely Huber-loss; see Table 4. As can be observed, the proposed loss function assists MIBINET in improving its performance substantially (both higher r and R^2) in all datasets. This signifies that our proposed loss is not dataset-specific but performs equally well over other datasets.

Next, we conduct Bland Altman analysis to show the agreement between predicted and ground truth IBI; see Fig. 10. In this analysis, we consider four different cases. The comparison between the true and predicted values is shown through the deviation plots and the Bland Altman plots of the IBI values and the corresponding BPM values. The plots in Fig. 10 show the mean and standard deviation in radar heart sound data. The LoA (limits of agreement) is lower in our method than that of state-of-the-art, which is shown in Figs. 10(b–c) and 10(k–l). The deviation plots also reveal a significant improvement over the HSMM method. We can clearly observe that there are some errors in the region of 0.1 s to 0.25 s for the HSMM deviation plot shown in 10a. Fig. 10(d–f) vs. Fig. 10(j–l) shows the clear distinction between Huber loss and proposed loss, respectively. As we can observe, the regressor with Huber loss has many errors beyond 0.05 s in the deviation plot. Moreover, the LoA and the mean are better for the regressor with the proposed loss.

4.5. Effectiveness of synthetic IBI generation

Again, we perform the Bland-Altman analysis to demonstrate the potency of our synthetic IBI generation process. It can be clearly noticed from Figs. 10(h–i) and 10(k–l) that our proposed augmentation method decreases the error and gives a better LoA. In purely quantitative terms, the RMSE is decreased by almost 3 ms for the Radar-HS dataset. Similarly, the results obtained through PCG data also show the effectiveness of this novel synthetic IBI augmentation. In PCG modality, we observe a performance improvement of around 10% in correlation coefficient and 40 ms in RMSE, as displayed in Table 5. Furthermore, the presence of a 20% decrease in R^2 without the inclusion of augmentation in the PCG modality indicates the occurrence of overfitting. The incorporation of augmentation techniques enhances the uniformity of prediction across all folds, hence indicating an improved level of generalizability.

4.6. Comparison with related models

Besides HSMM, we compare the performance of the MIBINET model against the current state-of-the-art lightweight models applicable to the Radar-HS dataset provided in [42]. To evaluate the performance of the neural network models alone, we test them without post-processing,

Table 6

Comparison with lightweight neural networks on the hold-out test set of Radar-HS (before post-processing filtering).

Model name	r (%)	R^2	Parameters ($\times 10^6$)
MobileNet V1 [52]	84.57	0.72	7.98
MobileNet V2 [53]	84.55	0.71	10.93
MobileNet V3 small [54]	84.44	0.72	1.28
MobileNet V3 large [54]	84.42	0.71	2.99
ResNet 34 [55]	78.06	0.63	7.07
ResNet 50 [55]	85.66	0.73	23.74
MIBINET	88.57	0.78	1.09

and the results are presented in Table 6. It is to be noted that all the peripheral settings were kept exactly the same for all the models to ensure a fair comparison. We notice that MIBINET outperforms the current state-of-the-art lightweight deep learning models by a margin of at least 3% in the correlation coefficient metric. Importantly, MIBINET, having the least number of parameters among all these models, claims itself as the best-suited real-time solution for contactless vital signs monitoring. It is crucial to emphasize that the inference time of MIBINET, within the confines of our current simulation setup, is a mere 0.42 ms.

5. Discussion and limitations

This work proposed a comprehensive DNN-based strategy (MIBINET) to surpass the present state-of-the-art (HSMM) in measuring IBI from mm-wave collected data. A key aspect of our work lies in the integration of novel signal processing schemes with CNN's excellent feature extraction capabilities. MIBINET is a unique lightweight architecture that features innovative post-processing techniques, including rolling window averaging and traditional filtering, which contribute to the improved performance of the model. The proposed approach achieved a cumulative 98.73% correlation coefficient, an elevated R^2 metric of 97%, and produced a meager 20.69 ms RMSE score over 11 different test subjects on Radar-HS data. This was a significant improvement over HSMM, which offered a 93.66% correlation coefficient, R-squared metric of 0.87%, and 47.94 ms RMSE.

The robustness and versatility of our proposed approach were further demonstrated by evaluating it on datasets containing ECG, PCG, and PPG signals, where its performance was found to be superior to HSMM in all of them. A custom-weighted regression loss was developed to train MIBINET, and a unique weight-saving approach based on a weighted metric was used. The novel weighted loss function improved the correlation coefficient by more than 1.4% and lowered the RMSE by more than 21 ms in Radar-HS data compared to the commonly used Huber-loss function. Additionally, a novel IBI augmentation technique was also introduced to negate the effects of dataset imbalance, which led to significant performance improvements.

This research is significant as it tackles the pervasive issue of data scarcity in physiological signal processing, enhancing the robustness of machine learning models in healthcare applications. By focusing on a broad spectrum of heart rate conditions, our study fills a crucial gap in the existing literature and offers actionable insights for clinicians and engineers alike.

Through the use of synthetic IBI generation augmentation, our model can robustly handle a wide array of heart rate values, showcasing its applicability in scenarios such as arrhythmia detection. Similarly, the rolling window-based pre-processing and augmentation technique has proven effective in diversifying datasets and overcoming dimensional constraints, thereby enhancing the model's generalization capability across various 1D signal processing tasks.

Theoretically, our research lays the groundwork for further explorations into the adaptability of neural network models in handling complex physiological signals. One theoretical observation is that derivative signals, which are shifted versions of basic signals, have the potential to enhance the dataset. Additionally, using weighted and mixed loss

functions has been found to enhance the optimization potential compared to using a single loss function. Practically, our techniques offer immediate value in healthcare applications requiring heart rate monitoring and arrhythmia detection, demonstrating promising pathways for advancements in remote patient monitoring systems.

While our study makes several significant contributions, it is not without its limitations. One of the main constraints pertains to the applicability of our synthetic IBI augmentation technique. Its effectiveness on non-stationary signals like ECG and PPG has yet to be thoroughly evaluated. Another concern is the model's performance in scenarios characterized by high Heart Rate Variability (HRV), where the robustness of MIBINET could potentially be put to the test.

In summary, although our model achieves high performance and versatility, the limitations outlined above suggest that further research is necessary to optimize its applicability across diverse clinical and real-world settings.

6. Conclusion and future works

In conclusion, our work has demonstrated the effectiveness of integrating novel signal processing techniques with CNN's feature extraction capabilities in estimating IBI values from various signal modalities such as radar-HS, PCG, ECG, and PPG. The proposed MIBINET approach showcased significant improvements over the current state-of-the-art methods. Our findings also highlighted the versatility and robustness of multifaceted MIBINET across diverse signal modalities, as its performance consistently surpassed that of the HSMM method. Moreover, our custom-weighted regression loss and the novel IBI augmentation technique effectively addressed dataset imbalance, leading to substantial performance enhancements. These results underline the potential of MIBINET in contactless vital signs monitoring, offering a promising solution for various real-life applications in medical, home, and transportation settings. Additionally, we believe that integrating our newly designed signal processing techniques can significantly enhance the performance of other 1D models.

Moving forward, there are several avenues for extending the capabilities of the MIBINET model. One immediate direction is to explore the use of raw radar data for end-to-end modeling. Such an exploration could offer increased accuracy and new insights. Additionally, the limitations of the synthetic IBI augmentation technique need to be addressed, particularly when applied to non-stationary signals like ECG and PPG. Furthermore, adapting MIBINET to multi-user scenarios will be crucial for making the model more universally applicable and practicable. Lastly, investigating the benefits of multi-modal approaches that integrate or fuse multiple types of signals promises to broaden the horizons of contactless vital signs monitoring. By addressing these future research challenges, we anticipate further refining and expanding MIBINET's existing capabilities, thereby amplifying its potential impact across a spectrum of real-life applications.

CRedit authorship contribution statement

Rafid Umayer Murshed: Conceptualization, Methodology, Software, Formal analysis, Writing – original draft, Validation, Writing – review & editing. **Md. Abrar Istiak:** Software, Validation, Data curation, Writing – original draft, Writing – review & editing, Investigation. **Md. Toufiqur Rahman:** Conceptualization, Visualization, Investigation, Data curation, Writing – original draft. **Zulqarnain Bin Ashraf:** Investigation, Data curation, Writing – original draft. **Md. Saheed Ullah:** Software, Writing – original draft. **Mohammad Saquib:** Writing – review & editing, Supervision, Resources, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

- [1] Thaulow E, Erikssen JE. How important is heart rate? *J Hypertens* 1991;9(Supplement 7):S27–30. <http://dx.doi.org/10.1097/00004872-199112007-00007>.
- [2] Centers for Disease Control and Prevention and National Center for Health Statistics, et al. About multiple cause of death, 1999–2019. CDC WONDER online database website. Centers for Disease Control and Prevention: Atlanta, GA, USA; 2019.
- [3] Tsao CW, Aday AW, Almarzooq ZI, Alonso A, Beaton AZ, Bittencourt MS, et al. Heart disease and stroke statistics—2022 update: a report from the American Heart Association. *Circulation* 2022;145(8):e153–639.
- [4] Volodina OV. Formation of future teachers' worldview culture by means of foreign-language education. *P Sci Edu* 2022;57(3):126–59. <https://pnojurnal.wordpress.com/2022/07/01/volodina-3/>.
- [5] Rajendra Acharya U, Paul Joseph K, Kannathal N, Lim CM, Suri JS. Heart rate variability: a review. *Med Biol Eng Comput* 2006;44(12):1031–51.
- [6] Schuster AK, Fischer JE, Thayer JF, Mauss D, Jarczok MN. Decreased heart rate variability correlates to increased cardiovascular risk. *Int J Cardiol* 2016;203:728–30.
- [7] Shaffer F, Ginsberg JP. An overview of heart rate variability metrics and norms. *Front Public Health* 2017;258.
- [8] Malik M. Heart rate variability: Standards of measurement, physiological interpretation, and clinical use: Task force of the European society of cardiology and the North American society for pacing and electrophysiology. *Ann Noninvasive Electrocardiol* 1996;1(2):151–81.
- [9] La Rovere MT, Bigger Jr JT, Marcus FI, Mortara A, Schwartz PJ, ATRAMI (Autonomic Tone and Reflexes After Myocardial Infarction) Investigators, et al. Baroreflex sensitivity and heart-rate variability in prediction of total cardiac mortality after myocardial infarction. *Lancet* 1998;351(9101):478–84.
- [10] Lan K-c, Raknim P, Kao W-F, Huang J-H. Toward hypertension prediction based on PPG-derived HRV signals: A feasibility study. *J Med Syst* 2018;42(6):1–7.
- [11] Harrigan RA, Chan TC, Brady WJ. Electrocardiographic electrode misplacement, misconnection, and artifact. *J Emerg Med* 2012;43(6):1038–44.
- [12] Bond RR, Finlay D, Nugent C, Breen C, Guldenring D, Daly M. The effects of electrode misplacement on clinicians' interpretation of the standard 12-lead electrocardiogram. *Eur J Intern Med* 2012;23(7):610–5.
- [13] Yang Z, Pathak PH, Zeng Y, Liran X, Mohapatra P. Monitoring vital signs using millimeter wave. In: *Proceedings of the 17th ACM international symposium on mobile Ad Hoc networking and computing*. 2016, p. 211–20.
- [14] Johnson JE, Shay O, Kim C, Liao C. Wearable millimeter-wave device for contactless measurement of arterial pulses. *IEEE Trans Biomed Circuits Syst* 2019;13(6):1525–34.
- [15] Töpfer F, Dudorov S, Oberhammer J. Millimeter-wave near-field probe designed for high-resolution skin cancer diagnosis. *IEEE Trans Microw Theory Tech* 2015;63(6):2050–9.
- [16] Di Meo S, Espin-Lopez PF, Martellosio A, Pasian M, Matrone G, Bozzi M, et al. On the feasibility of breast cancer imaging systems at millimeter-waves frequencies. *IEEE Trans Microw Theory Tech* 2017;65(5):1795–806.
- [17] Hoshi N, Nikawa Y, Kawai K, Ebisu S. Application of microwaves and millimeter waves for the characterization of teeth for dental diagnosis and treatment. *IEEE Trans Microw Theory Tech* 1998;46(6):834–8.
- [18] Lin F, Hu W, Li A. Millimeter-wave technology for medical applications. In: *2012 IEEE MTT-S international microwave workshop series on millimeter wave wireless technology and applications*. 2012, p. 1. <http://dx.doi.org/10.1109/IMWS2.2012.6338252>.
- [19] Alizadeh M, Shaker G, Almeida JCMD, Morita PP, Safavi-Naeini S. Remote monitoring of human vital signs using mm-Wave FMCW radar. *IEEE Access* 2019;7:54958–68. <http://dx.doi.org/10.1109/ACCESS.2019.2912956>.
- [20] J. Bancifra J, Tarlac State University, Tarlac City, Philippines. Supervisory practices of department heads and teachers' performance: Towards a proposed enhancement program. *APJAET - J Asia Pacific J Adv Educ Technol* 2022;25–33. <https://orcid.org/0000-0003-0641-1305>.
- [21] Wang C-C, Trivedi SB, Jin F, Stepanov S, Chen Z, Khurgin J, et al. Human life signs detection using high-sensitivity pulsed laser vibrometer. *IEEE Sens J* 2007;7(9):1370–6. <http://dx.doi.org/10.1109/JSEN.2007.905041>.
- [22] Wang F, Zeng X, Wu C, Wang B, Liu KR. mmhrv: Contactless heart rate variability monitoring using millimeter-wave radio. *IEEE Internet Things J* 2021;8(22):16623–36.
- [23] Zhang J, Chen Y, Chen T, et al. Health-radio: towards contactless myocardial infarction detection using radio signals. *IEEE Trans Mob Comput* 2020.
- [24] Ahmad A, Roh JC, Wang D, Dubey A. Vital signs monitoring of multiple people using a FMCW millimeter-wave sensor. In: *2018 IEEE radar conference*. 2018, p. 1450–5. <http://dx.doi.org/10.1109/RADAR.2018.8378778>.
- [25] Schires E, Georgiou P, Lande TS. Vital sign monitoring through the back using an UWB impulse radar with body coupled antennas. *IEEE Trans Biomed Circuits Syst* 2018;12(2):292–302. <http://dx.doi.org/10.1109/TBCAS.2018.2799322>.

- [26] Wu S, Sakamoto T, Oishi K, Sato T, Inoue K, Fukuda T, et al. Person-specific heart rate estimation with ultra-wideband radar using convolutional neural networks. *IEEE Access* 2019;7:168484–94. <http://dx.doi.org/10.1109/ACCESS.2019.2954294>.
- [27] Wang P, Qi F, Liu M, Liang F, Xue H, Zhang Y, et al. Noncontact heart rate measurement based on an improved convolutional sparse coding method using IR-UWB radar. *IEEE Access* 2019;7:158492–502. <http://dx.doi.org/10.1109/ACCESS.2019.2950423>.
- [28] Vinci G, Lindner S, Barbon F, Mann S, Hofmann M, Duda A, et al. Six-port radar sensor for remote respiration rate and heartbeat vital-sign monitoring. *IEEE Trans Microw Theory Tech* 2013;61:2093–100. <http://dx.doi.org/10.1109/TMTT.2013.2247055>.
- [29] Li C, Ling J, Li J, Lin J. Accurate Doppler radar noncontact vital sign detection using the RELAX algorithm. *IEEE Trans Instrum Meas* 2010;59(3):687–95. <http://dx.doi.org/10.1109/TIM.2009.2025986>.
- [30] Petrović VL, Janković MM, Lupšić AV, Mihajlović VR, Popović-Božović JS. High-accuracy real-time monitoring of heart rate variability using 24 GHz continuous-wave Doppler radar. *IEEE Access* 2019;7:74721–33. <http://dx.doi.org/10.1109/ACCESS.2019.2921240>.
- [31] Sakamoto T, Imasaka R, Taki H, Sato T, Yoshioka M, Inoue K, et al. Feature-based correlation and topological similarity for interbeat interval estimation using ultrawideband radar. *IEEE Trans Biomed Eng* 2015;63(4):747–57.
- [32] Immervoy I, Fedotov P. Ultra wideband radar systems: advantages and disadvantages. In: 2002 IEEE conference on ultra wideband systems and technologies (IEEE Cat. No.02EX580). 2002, p. 201–5. <http://dx.doi.org/10.1109/UWBST.2002.1006348>.
- [33] Li M, Lin J. Wavelet-transform-based data-length-variation technique for fast heart rate detection using 5.8-GHz CW Doppler radar. *IEEE Trans Microw Theory Tech* 2018;66(1):568–76. <http://dx.doi.org/10.1109/TMTT.2017.2730182>.
- [34] Brüser C, Stadthanner K, Brauers A, Leonhardt S. Applying machine learning to detect individual heart beats in ballistocardiograms. In: 2010 Annual international conference of the IEEE engineering in medicine and biology. IEEE; 2010, p. 1926–9.
- [35] Malešević N, Petrović V, Belić M, Antfolk C, Mihajlović V, Janković M. Contactless real-time heartbeat detection via 24 GHz continuous-wave Doppler radar using artificial neural networks. *Sensors* 2020;20(8). <http://dx.doi.org/10.3390/s20082351>, URL <https://www.mdpi.com/1424-8220/20/8/2351>.
- [36] Biswas D, Everson L, Liu M, Panwar M, Verhoeef B-E, Patki S, et al. CorNET: Deep learning framework for PPG-based heart rate estimation and biometric identification in ambulant environment. *IEEE Trans Biomed Circuits Syst* 2019;13(2):282–91. <http://dx.doi.org/10.1109/TBCAS.2019.2892297>.
- [37] Chang X, Li G, Xing G, Zhu K, Tu L. DeepHeart: A deep learning approach for accurate heart rate estimation from PPG signals. *ACM Trans Sensor Netw* 2021;17(2):1–18.
- [38] Huang B, Lin C-L, Chen W, Juang C-F, Wu X. A novel one-stage framework for visual pulse rate estimation using deep neural networks. *Biomed Signal Process Control* 2021;66:102387.
- [39] Katsaouni N, Aul F, Kriskcher L, Schmalhofer S, Hedrich L, Schulz MH. Energy efficient convolutional neural networks for arrhythmia detection. *Array* 2022;13:100127.
- [40] Cheng J, Yue B, Song R, Liu Y, Li C, Chen X. Motion-robust anterior-posterior imaging ballistocardiography for non-contact heart rate measurements. *Biomed Signal Process Control* 2023;86:105307. <http://dx.doi.org/10.1016/j.bspc.2023.105307>, URL <https://www.sciencedirect.com/science/article/pii/S1746809423007930>.
- [41] Zhang H, Jian P, Yao Y, Liu C, Wang P, Chen X, et al. Radar-beat: Contactless beat-by-beat heart rate monitoring for life scenes. *Biomed Signal Process Control* 2023;86:105360. <http://dx.doi.org/10.1016/j.bspc.2023.105360>, URL <https://www.sciencedirect.com/science/article/pii/S1746809423007930>.
- [42] Shi K, Schellenberger S, Will C, Steigleder T, Michler F, Fuchs J, et al. A dataset of radar-recorded heart sounds and vital signs including synchronised reference sensor signals. *Sci Data* 2020;7(1):1–12.
- [43] Jana GC, Sharma R, Agrawal A. A 1D-CNN-spectrogram based approach for seizure detection from EEG signal. *Procedia Comput Sci* 2020;167:403–12.
- [44] Panwar M, Gautam A, Biswas D, Acharyya A. PP-net: A deep learning framework for PPG-based blood pressure and heart rate estimation. *IEEE Sens J* 2020;20(17):10000–11.
- [45] Ramachandran P, Zoph B, Le Q. Swish: a Self-Gated Activation Function. 2017.
- [46] Huber PJ. Robust estimation of a location parameter. *Breakthr Stat Methodol Distribution* 1992;492–518.
- [47] Kingma DP, Ba J. Adam: A method for stochastic optimization. 2014, arXiv preprint [arXiv:1412.6980](https://arxiv.org/abs/1412.6980).
- [48] Springer DB, Tarassenko L, Clifford GD. Logistic regression-HSMM-based heart sound segmentation. *IEEE Trans Biomed Eng* 2015;63(4):822–32.
- [49] Wagner P, Strodthoff N, Bousselet R-D, Kreiseler D, Lunze FI, Samek W, et al. PTB-XL, a large publicly available electrocardiography dataset. *Sci Data* 2020;7(1):1–15.
- [50] Milivojević MS, Gavrovska A, Reljin IS, et al. Python based physiological signal processing for vital signs monitoring. 2017.
- [51] Jarchi D, Casson AJ. Description of a database containing wrist PPG signals recorded during physical exercise with both accelerometer and gyroscope measures of motion. *Data* 2016;2(1):1.
- [52] Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications. 2017, arXiv preprint [arXiv:1704.04861](https://arxiv.org/abs/1704.04861).
- [53] Sandler M, Howard A, Zhu M, Zhmoginov A, Chen L-C. Mobilenetv2: Inverted residuals and linear bottlenecks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2018, p. 4510–20.
- [54] Howard A, Sandler M, Chu G, Chen L-C, Chen B, Tan M, et al. Searching for mobilenetv3. In: Proceedings of the IEEE/CVF international conference on computer vision. 2019, p. 1314–24.
- [55] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2016, p. 770–8.