Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2019.DOI

Person-specific Heart Rate Estimation with Ultra-wideband Radar Using Convolutional Neural Networks

SHUQIONG WU[1], TAKUYA SAKAMOTO[2], (Senior Member, IEEE), KENTARO OISHI[1], TORU SATO[3], (Member, IEEE), KENICHI INOUE[4], TAKESHI FUKUDA[4], KENJI MIZUTANI[4], and HIROYUKI SAKAI[5]

[1]Graduate School of Informatics, Kyoto University, Yoshida-Honmachi, Sakyo Ward, Kyoto 606-8501, Japan
[2]Graduate School of Engineering, Kyoto University, Kyotodaigaku-Katsura, Nishikyo-ku, Kyoto 615-8510, Japan
[3]Institute for Liberal Arts and Sciences, Kyoto University, Yoshida-Honmachi, Sakyo Ward, Kyoto 606-8501, Japan
[4]Institute for Sensors and Devices, Technology Innovation Division, Panasonic Corporation, 3-1-1, Yakumo-Nakamachi, Moriguchi, Osaka, Japan
[5]Technology Liaison Department, Innovation Strategy Office, Panasonic Corporation, Kadoma 1006, Osaka, Japan

Corresponding author: Shuqiong Wu (e-mail: wusq@i.kyoto-u.ac.jp).

This study was supported in part by JSPS KAKENHI Grant Numbers 15K18077, 15KK0243, and 19H02155, the Japan Science and Technology Agency PRESTO Grant Number JPMJPR1873, and JST COI Program JPMJCE1307.

ABSTRACT Vital-sign estimation using ultra-wideband (UWB) radar is preferable because it is contactless and less privacy-invasive. Recently, many approaches have been proposed for estimating heart rate from UWB radar data. However, their performance is still not reliable enough for practical applications. To improve the accuracy, this study employs convolutional neural networks to learn the special patterns of the heartbeats. In the proposed system, skin displacements of the target person are measured using UWB radar, and the radar signal is converted to a two-dimensional matrix, which is used as the input of the designed neural networks. Meanwhile, two triangular waves corresponding to the peaks and valleys in an electrocardiogram are adopted as the output of the networks. The proposed system then identifies each individual and estimates the heart rate automatically based on the already trained neural networks. The estimation error of the interbeat interval computed using our approach was reduced to 4.5 ms in the best case; and 48.5 ms in the worst case. Experiment results show that the proposed approach significantly outperforms a conventional method. The proposed machine learning approach achieves both personal identification and heart rate estimation simultaneously using UWB radar data for the first time. Moreover, this study found that using the respiration and heartbeat components together may enhance the accuracy of heart rate estimation, which is counter-intuitive, because the respiration is usually believed to interfere with the heartbeat.

INDEX TERMS Ultra-wideband radar, heart rate, vital signs, convolutional neural networks

I. INTRODUCTION

In recent decades, computer-based vital signs monitoring has played an increasing role in the medical care and nursing fields. For example, in hospitals, child care centers, and nursing homes, heart rate monitoring can avoid many accidents caused by cardiovascular diseases. Considering that multiple persons commonly share the same space, an ideal monitoring system would provide not only accurate vital data, but also a function to identify each individual. In this kind of person-specific long-term monitoring system, a correspondence between the data and the subject should be built because a mismatch may result in serious problems. Monitoring approaches using wearable or contact devices are still the mainstream because they are reliable; and do not require any extra processing of personal identification. For example, electrocardiography (ECG) and photoplethysmography (PPG) sensors are usually used in heartbeat monitoring [1]–[3]. Sensors worn on the wrist [4], [5] or in the ear [6] have also been designed to monitor vital signs. However, wearing these devices is troublesome; and may cause skin discomfort [7]. Some researchers have designed contactless monitoring using cameras [8]–[11], but these methods are sensitive to illumination change and visual occlusion.

In contrast to these sensor-based or camera-based approaches, radar has been considered a more comfortable and reliable option for monitoring vital signs [12]–[17]. To

estimate the heartbeat and respiration rate, some approaches use the Fourier transform [12]–[14] and others adopt feature extraction [15]–[17]. Recently, ultra-wideband (UWB) radar has been proved to be more efficient than continuous-wave (CW) radar in measuring small skin displacement because it can suppress interfering echoes from different distances [19]–[30]. Nevertheless, radar-based monitoring systems have two major problems. One is that the performance is always degraded by the interference between the heartbeat and respiration components. To address this issue, some researchers have adopted filtering based on Fourier analysis, wavelet analysis, eigen features, and other methods [21], [22], [38]. Although there has been good progress on this front, it is still challenging to separate the respiration and heartbeat components [18], [23]. The other challenge is regarding the identification of individuals among multiple persons, because, in contrast to the camera-based methods, identifying individuals requires extra processing when using a radar for the measurements.

To solve these problems, the present paper applies convolutional neural networks (CNNs) instead of traditional techniques to our heartbeat monitoring system. Although some recent studies have reported that a CNN is efficient in classifying ECG data [31]–[36], most such studies aim to classify normal and abnormal ECG waveforms without any radar signals involved. To the authors' knowledge, no study has been conducted for processing radar data to identify individuals as well as estimate the heart rate simultaneously. The heart rate is estimated by measuring the interval between two adjacent unique patterns of the heartbeat. Therefore, we exploit CNNs to regress a time-varying curve which carries the unique patterns.

In this paper, a time series of the phase of radar signals is used as the input of the CNNs, and two triangular waves that correspond to the R and S waves respectively are designed to be the output of the CNNs, which is trained by solving a regression problem. The contributions of the present paper are listed as follows: (1) This paper proposes a novel heartbeat monitoring system that significantly improves estimation accuracy. (2) Instead of using the raw ECG waveforms, this paper introduces two triangular patterns for training the CNNs. Triangular patterns corresponding to the R and S waves (early and late ventricular depolarizations) are proposed to achieve a better regression than that realized by using ECG directly. (3) This paper has found for the first time that the respiration component of the radar data improves the accuracy of the heartbeat estimation, which defies the common belief in this field that the respiration component interferes with the heartbeat component and degrades the overall accuracy. (4) This paper realizes person-specific heartbeat monitoring using a radar system for the first time, because the identification of each subject can be automatically achieved by the CNNs.

When a CNN is used for person-independent classification, a large number of training data are required, and the training process needs an unacceptably large amount of computational resources. However, in our system, we can train each CNN specifically for each subject in advance; so that it can automatically identify each subject and estimate his or her heart rate. Our study in this paper found that a small number of training data were sufficient to achieve high performance in person-specific monitoring. The proposed approach is expected to play a crucial role in realizing an accurate long-term and noncontact heartbeat monitoring system in a variety of applications.

II. HEARTBEAT MONITORING USING UWB RADAR

This section introduces the basic theory of heartbeat monitoring using UWB radar data, and also explains an existing algorithm that outperforms many other methods [18].

A. DISPLACEMENT MEASUREMENT BASED HEARTBEAT ESTIMATION

In this subsection, we briefly explain the basic principle of heartbeat estimation using a UWB radar system. Detailed analyses can be found in [19], [20]. We define the received radar signal at time t and range r as $s(t, r)$, which is complex-valued, where the real and imaginary parts are the in-phase and quadrature components. We manually select range r_0 corresponding to the position of the target subject. Then, a displacement of the target can be obtained as

$$d(t) = \text{unwrap}\left\{\angle s(t, r_0)\right\} \lambda/4\pi, \tag{1}$$

where $d(t)$ is the estimated displacement of time t, $\angle$ is the phase operator of a complex number and obtained using an arctangent demodulation, λ is the wavelength at the center frequency, and unwrap $\{\cdot\}$ is an unwrapping process that obtains a smooth phase sequence considering the phase ambiguity of $2n\pi$, where n is an integer. The unwrap function used in this study corrects the phase of a complex-valued sequence by adding multiples of $\pm 2\pi$ when absolute jumps between adjacent samples are greater than or equal to π. Therefore, the resultant phase sequence becomes smooth [42]. We note that the estimated displacement contains both respiration and heartbeat components.

B. TOPOLOGY-FEATURE-BASED ALGORITHM

A topology-feature-based method (TF Method) for the radar-based heartbeat estimation has been shown to be one of the most accurate existing algorithms [18]. The skin displacements corresponding to the heartbeats have quasi-periodic patterns, so the computation of the interbeat interval (IBI; also the reciprocal of the instantaneous heart rate) can be realized by measuring the intervals between adjacent quasi-periodic waveforms that are topologically similar. This kind of topological similarity is exploited to improve the accuracy in the measurement of heartbeat using UWB radar systems [18]. However, the TF Method still has some challenges, e.g., the performance is occasionally degraded by the interference between the heartbeat and respiration components. Moreover, individuals cannot be identified based on radar signals

using the TF Method; that needs to be done by other techniques. Because no other existing techniques can solve these problems, we explore the use of machine learning methods to replace the conventional signal processing methods. The major purpose of the research is to verify the efficiency of machine learning methods when they are applied to radar-based heartbeat monitoring and individual identification. Among the various machine learning algorithms, we choose CNNs because they have been reported to be remarkably effective in processing various types of data [43].

III. PROPOSED ALGORITHM

A. USER SCENARIO FOR THE PROPOSED ALGORITHM

In this section, first we describe a user scenario for the proposed heartbeat monitoring system. The proposed system is designed to be used in a space shared by several people, such as a home, office, or hospital. In the initialization setup, a new user is required to register his/her name and user ID, and remain seated or lying still for 100 s a few meters away from the radar system, with ECG electrodes attached to his/her chest. After the measurement, the pair of radar and ECG data are used to train the user's person-specific neural network. Once this initialization setup is complete, the user can always be identified and his/her heart rate can be monitored when the user is in static scenarios. Note that heart rate estimation is person-specific in the proposed system. If the CNNs are trained for general people (i.e., not person-specific), then the system cannot identify which user is under test, which makes heartbeat monitoring useless in a space shared by multiple people.

B. STRUCTURE DESIGN OF THE CNN

Fig. 1 shows the layers of the CNN used in this study. It contains one input layer, two convolutional layers, one rectified linear unit (ReLU) layer, one fully connected layer, and one regression layer. The size and number of filters are determined empirically as shown in Fig. 1. The CNN in this study do not contain pooling layers because a study [39] reported that the use of pooling layers can degrade the

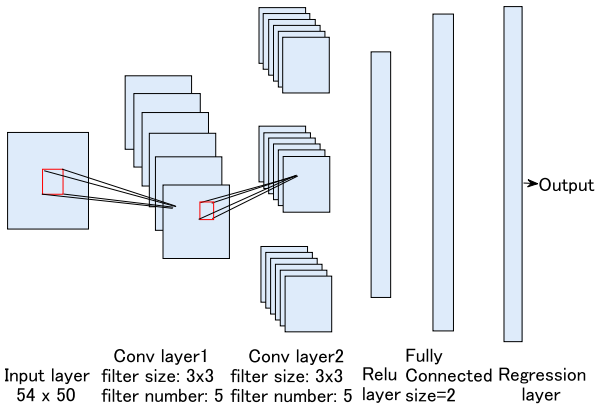


FIGURE 1. Layers adopted for each CNN: the five CNNs share the same layer layout, but they are trained independently and in parallel.

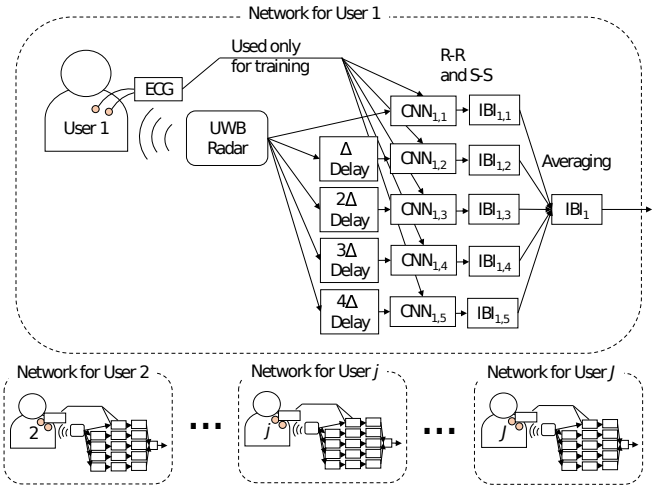


FIGURE 2. Proposed networks that contain five independent CNNs. Each network is trained using radar and ECG data for each user.

accuracy of estimation. As illustrated in Fig. 1, the input sequence of each CNN is a two-dimensional (2D) matrix, and the output of each CNN consists of two one-dimensional time sequences of signals. The regression layer here aims at changing a classification network into a regression network. The CNN is trained to find a connection between a time series of ECG data and a time series of radar data.

We specifically selected the CNN rather than other machine learning algorithms because the CNN was reported to be more effective in learning nonlinear relationships than other machine-learning methods [44]. Additionally, estimating the heart rate using radar data is considered highly nonlinear. The main focus of this study is to demonstrate the effectiveness of the machine learning approach in radar-based heartbeat monitoring. It is also important to compare different machine learning algorithms for radar-based heartbeat monitoring, which will be performed in future studies.

C. INPUT SAMPLING FROM UWB RADAR DATA

If the radar and ECG devices share the same clock and they are completely synchronized, a single CNN is sufficient for the estimation of the heart IBI of each user. However, when radar and ECG devices have independent local clocks, the devices cannot be always completely synchronized, which is the case in our experiment; a time deviation between the devices degrades the accuracy of the heart IBI estimation when using a single CNN. Because the time deviation between the devices varies over time, an initial calibration cannot resolve the synchronization issue. This is why we introduced a parallel structure with multiple CNNs, as shown in Fig. 2. Specifically, the proposed network contains five CNNs, where the input sequences to these CNNs are time-shifted by zero, Δt, $2\Delta t$, $3\Delta t$, and $4\Delta t$, where $\Delta t = 0.12$ s. The output sequences from the CNNs are averaged to estimate the heart IBI, which enables the proposed CNN-based algorithm to tolerate a time delay of up to 0.48 s. Increasing the number of CNNs in the proposed network

would make the system even more robust against a longer time deviation. The selection of the number of CNNs and time shift intervals depends on the actual synchronization accuracy of the devices.

In our system, the time sampling interval of the radar is $\delta = 0.24$ ms and that of the ECG signal is 2.0 ms. Radar displacement signal $d(t)$ is filtered using a bandpass filter to improve the SNR. Then, the signal is normalized so that the displacement signal values fall into the range [0,1]. Next, N samples of the radar displacement signal are extracted to form an input sequence, where $N = 27,000$, which corresponds to 6.43 s, was empirically selected. The input sequence is thus expressed as $\phi(t) = \{d(t - T/2 + \delta), d(t-T/2+2\delta), \cdots, d(t+T/2)\}$. Its time-shifted versions $\phi(t - 2\Delta t)$, $\phi(t - \Delta t)$, $\phi(t)$, $\phi(t + \Delta t)$, and $\phi(t + 2\Delta t)$ are the input sequences for the five CNNs, where $\Delta t = 0.12$ s. The input sequences with a length of 2,700 (the average of every 10 points) are reshaped into 54×50 matrices.

D. OUTPUT DATA TRANSFORMATION BASED ON TRIANGULAR WAVES

To estimate the heart rate, the output waveform of each CNN is designed to be related to the ECG waveform so that the heart rate can be computed from the output. A straightforward approach is to use the raw ECG waveform for training the CNNs, as shown in Fig. 3. The ECG waveform, however, changes sharply during the RST interval, which is different from the smooth radar waveform associated with the mechanical activity of a human heart. This difference makes regression a challenging task. To solve this problem, we use two triangular waveforms corresponding to the R and S waves respectively for training instead of the original ECG waveform. Thus, the regression layer of each CNN is used to estimate the two triangular waveforms.

For each R–R interval, we use the triangular wave shown in Fig. 4(a) to replace the ECG waveform. Similarly, we use the triangular wave shown in Fig. 4(b) to replace the ECG waveform in each S–S interval. Because the R and S waves are very close, we use an inverted triangular waveform to represent the S wave to make the waves easier to visualize. Consequently, the sharply changing RST interval of the ECG

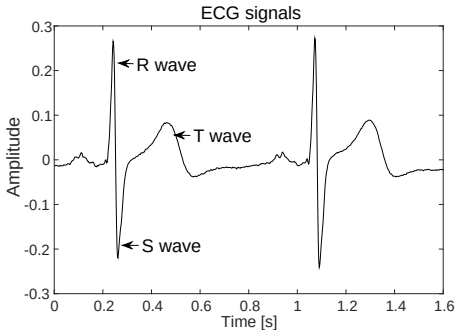
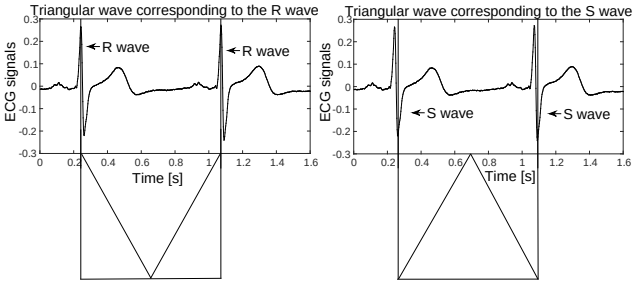


FIGURE 3. Example of ECG signals (the R, S, and T waves are specified).



(a) Replace the R–R interval signals (b) Replace the S–S interval signals

FIGURE 4. Transformation of the ECG waveforms into two triangular waves.

is transformed into two gradually changing triangular waves. Next, we give an example to illustrate the relationship among the radar signals, ECG signals, transformed triangular waves, and output of the trained CNNs.

First, we trained the CNNs directly using the ECG waveforms for comparison. Fig. 5 shows the radar signal, ECG signal, and estimated curve from the trained CNNs on the same timeline. Here, the CNNs are trained by other data (not the data in Fig. 5) from the same participant. We find that the output of the CNNs is totally different from the ECG waveform even we use the ECG as the training data. This is because the sharply changing RST waveforms are difficult to determine using UWB radar data alone. In the last row of Fig. 5, the peaks and valleys corresponding to the heartbeat appear after 19,500 iterations, but the output is not similar to the ECG waveform. It is similar to a triangular waveform. This gave us an idea that triangular waveforms should be used instead of the ECG signals for training the CNNs.

Fig. 6 shows the results when using triangular waveforms, where the correspondence of the radar, ECG, triangular waves created from the ECG, and the output of the CNNs trained by the triangular waveforms, is illustrated. The training and test data used in Fig. 6 are the same as those used in

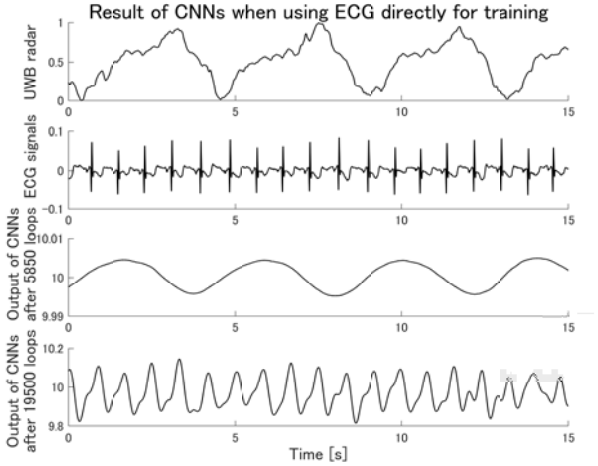


FIGURE 5. Relationship among the radar, ECG, and output of the CNNs.

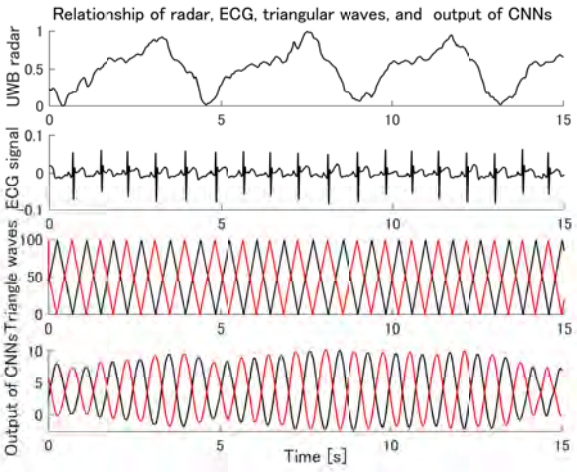


FIGURE 6. Relationship among the radar, ECG, triangular waves, and outputs of the CNNs: the four types of data are on the same timeline. The black curves of the first and the second rows show the radar and ECG signals, respectively. In the third row, the red curve denotes the triangular waveform corresponds to the R wave shown in Fig.4(a), and the black curve denotes the triangular wave corresponds to the S wave shown in Fig.4(b). In the last row, the red and black curves show the estimated triangular waves related to the R and S waves, respectively, and the two curves are the outputs of the CNNs.

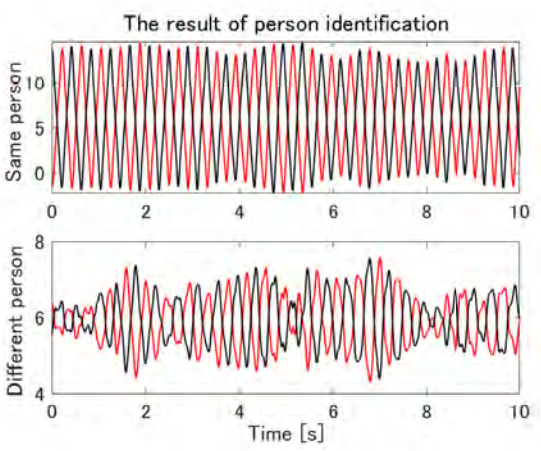


FIGURE 7. Personal identification: The upper row shows the outputs of the CNNs when the training and test data are from the same subject. The lower row shows the outputs of the CNNs when the training and test data are from different subjects.

Fig. 5. Fig. 6 shows that the proposed CNN-based algorithm output waveforms (in the 4th row) similar to the triangular waveforms (in the 3rd row), which resulted in a heart IBI error of 26.0 ms, whereas the error was 84.0 ms when the CNNs were trained using the raw ECG signal instead of the triangular waveforms. Moreover, when using the triangular waveform, the number of iterations required for training was 5,850, whereas it was 19,500 (shown in Fig. 5) when using the raw ECG signal, which indicates that the proposed algorithm using the triangular waveforms not only improves accuracy, but also saves training time.

Next, we investigate the performance of the proposed CNN-based algorithm using a triangular waveform that corresponds to the R wave alone and S wave alone instead of using both the R and S waves. The average IBI error over seven participants when using both the R and S waves, R wave alone, and S wave alone were 20.6 ms, 27.1 ms, and 30.8 ms, respectively. This accuracy improvement can be explained by the increase in the amount of information used for training the CNNs. The R-R interval and S-S interval are not always the same. Thus, using both the R and S waves allows us to exploit more information about the heartbeat of the participants.

E. PERSON IDENTIFICATION BASED ON CNNS

The proposed approach identifies individuals based on the classification capability of CNNs. We suppose that the number of users is J, we train J networks, where the j-th network is trained using the radar and ECG data from the j-th user, as shown in Fig. 2. After training, the CNNs are used for both personal identification and heart IBI estimation. When the radar signal from the j-th user is fed to the J trained networks, the j-th network is most likely to generate a triangular signal with an almost constant amplitude, whereas the other networks generate relatively weak signals with amplitude fluctuations, as shown in Fig. 7. The proposed CNN-based algorithm detects the constant-amplitude triangular signal using the average peak values defined as $\sum_{i=1}^{N_\mathrm{p}} V_i/N_\mathrm{p}$, where V_i is the i-th peak value and N_p is the number of peaks. The proposed CNN-based algorithm identifies individuals and estimates the heart IBI simultaneously as follows:

Step 1: Measure the radar and corresponding ECG signals from the j-th user $(j = 1, \cdots, J)$ and train the j-th network using the pair of data.

Step 2: After the k-th user is measured using the UWB radar, input the obtained radar signal to all J networks sequentially. Then compare the J output waveforms from these networks and select the network with the largest average peak value as a match.

Step 3: We suppose that the r-th network was selected in Step 2. This mean that the k-th user is recognized by the r-th network. The output from the r-th network will be used to estimate the heart IBI of the k-th user.

When k and r are matched as the above method shows, the proposed CNN-based algorithm is expected to generate accurate estimates of the heart IBI. In this study, seven participants were tested, and all of them were correctly identified. Thus, once our CNNs are trained person-specifically, each subject can be identified using radar data even when the system processes the data from multiple subjects simultaneously. After personal identification, the matched CNN can be used to estimate the heart IBI of the same subject. This system is especially suitable for a scenario in which one space is shared by several people, and their vital signs are

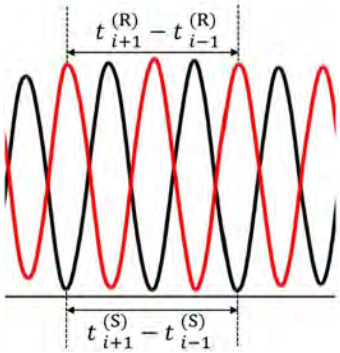


FIGURE 8. Triangular waves $W_1(t)$ and $W_2(t)$, represented by the red and black curves. Then the IBI are computed by $(t^{(R)}_{i+1} - t^{(R)}_{i-1} + t^{(S)}_{i+1} - t^{(S)}_{i-1})/4$, where $t^{(R)}_i$ and $t^{(S)}_i$ are the peak and valley positions of $W_1(t)$ and $W_2(t)$, respectively.

intermittently monitored and recorded over a long period of time in healthcare applications.

F. CALCULATION OF IBI BASED ON THE CNN OUTPUT

As discussed above, the proposed CNNs produce two triangular waveforms $W_1(t)$ and $W_2(t)$ that correspond to the R and S waves of the ECG. Fig. 8 shows how the IBI (the reciprocal of the heart rate) values are calculated from the two triangular signals. In Fig. 8, the red and black lines represent $W_1(t)$ and $W_2(t)$, respectively. The peaks of $W_1(t)$ and valleys of $W_2(t)$ are detected as $t^{(R)}_i$ and $t^{(S)}_i$ by the following equations:

$$t^{(R)}_i = t, \text{ if } dW_1(t)/dt = 0 \text{ and } d^2W_1(t)/dt^2 < 0. \quad (2)$$

$$t^{(S)}_i = t, \text{ if } dW_2(t)/dt = 0 \text{ and } d^2W_2(t)/dt^2 > 0. \quad (3)$$

Then, the i-th IBI is estimated by averaging $(t^{(R)}_{i+1} - t^{(R)}_{i-1} + t^{(S)}_{i+1} - t^{(S)}_{i-1})/4$, where averaging can suppress the influence of random components.

IV. EXPERIMENTAL RESULTS

In this section, first we explain the system settings in the experiments, and then we show the comparisons of the conventional and proposed techniques.

A. SYSTEM MODEL AND EXPERIMENTAL SETTINGS

We used a 79-GHz UWB multiple-input multiple output (MIMO) radar system with four transmitting antennas and four receiving antennas. The radar chips were developed by Imec (Heverlee, Belgium), and we integrated the two Imec chips with a digital control and interface circuit on the radio frequency circuit board. Using this system, a total of 16 radar signal channels were collected. UWB radar in the 79-GHz band has a high-range resolution because the wide bandwidth of 4.0 GHz is available, and its cost has recently become significantly inexpensive. Additionally, 79-GHz radar has a short wavelength, and thus is more sensitive to vital signs. Considering all these merits, we chose 79-GHz radar in this study. The pitch between any two adjacent antennas is 4.6 mm, which corresponds to 0.92 wavelength [37], [38]. We measured the radar echo reflected from the shoulder of each participant. The radar system was placed 0.9 m from the floor and 1.1 m from the shoulder. We synthesized the data from 16 channels and applied a beamforming technique to maximize the SNR. The beamforming uses the eigenvector that corresponds to the maximum eigenvalue of the correlation matrix of the signal vector as the weighting vector [38]. The experimental setup is shown in Fig. 9, where the radar measures from the left-hand side of the seated person. The radar echo from a range that corresponds to the left shoulder is extracted and used for personal identification and heartbeat monitoring in this study.

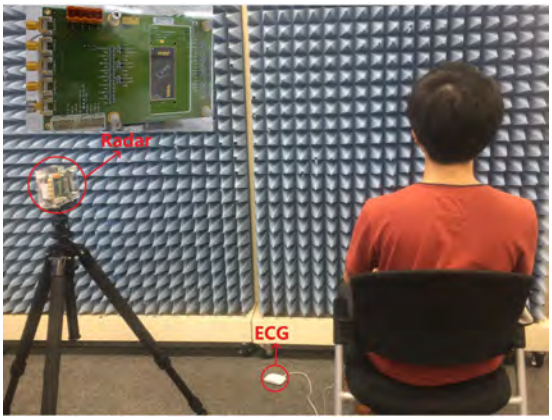


FIGURE 9. Experimental environment: The radar is placed on the left side of the participant. The distance between the radar and the shoulder of the participant is 1.1 m.

The ECG signals were used as the ground truth in the evaluation. We used an ECG device (ECG15102017, PLUX Wireless Biosignals S.A., Arruda dos Vinhos, Portugal), whose three electrodes were attached to the upper chest of each participant. The sampling frequency and resolution of the device were 500 Hz and 16 bits, and the data were wirelessly transmitted to a receiver via Bluetooth. The ECG receiver was not synchronized with the radar system.

We measured UWB radar and ECG signals for seven participants (P1 to P7), whose age and body mass index (BMI) are listed in Table I. The participants were requested to be seated and breathing normally. The measurement for each participant took 300 s. The training procedure is detailed in Section III. For each participant, we used 50,000 samples (100 s) for training and another 25,000 samples (50 s) for testing. The training for each participant took approximately 10 min using the Parallel Computing Toolbox (GPU) of MATLAB. The CPU and memory of the desktop we used are i7-6700 @ 3.40 GHZ 3.41 GHZ and 32 GB.

B. COMPARISONS AND RESULTS

We compare the proposed approach with the TF Method [18] for calculating IBI. For the TF Method, to extract the heartbeat component while rejecting the respiration component, a

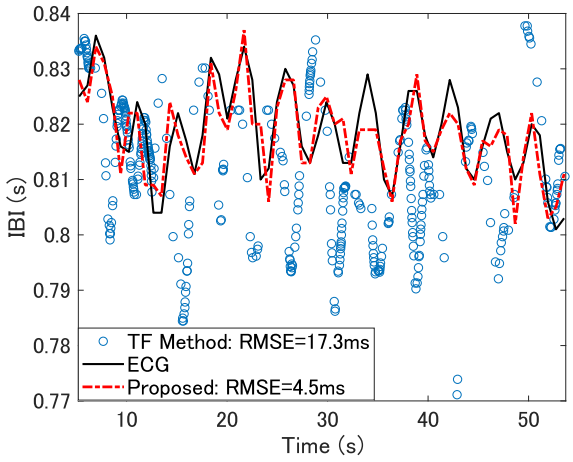


FIGURE 10. TF Method vs. the proposed method for estimating IBIs for P1.

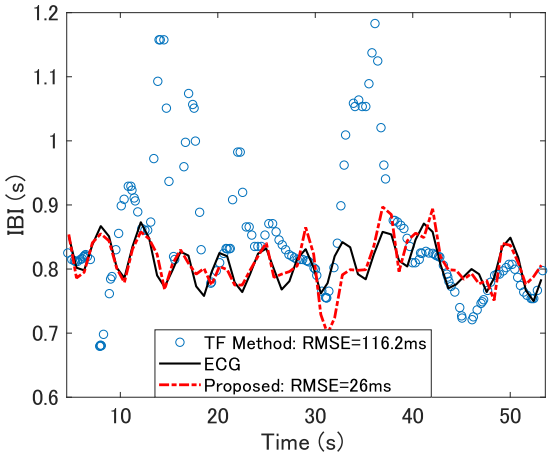


FIGURE 11. TF Method vs. the proposed method for estimating IBIs for P2.

bandpass filter with low/high cut-off frequencies of 3.1 and 12.0 Hz was used. Because the respiration component was much larger than the heartbeat component, the high-order harmonics of the respiration components could interfere with the heartbeat component. This is why we set the low cut-off frequency of the bandpass filter to be higher than that in other studies. Note that the proposed CNN-based algorithm does not require filtering, as detailed in Section V-C. The root mean square error (RMSE), and the Mean±STD of the absolute error are used for the evaluation below.

Figs. 10 to 12 show three examples of the comparison. In these figures, the horizontal axis represents the time, and the vertical axis represents the IBIs. Here, the black curves show the IBIs estimated using the ECG signals, the blue circles show the IBIs estimated using the TF Method, and the red curves show the IBIs estimated using the proposed method. These figures show that the proposed method is more accurate than the TF Method. From these figures, we also see that the IBIs estimated using the TF Method are occasionally

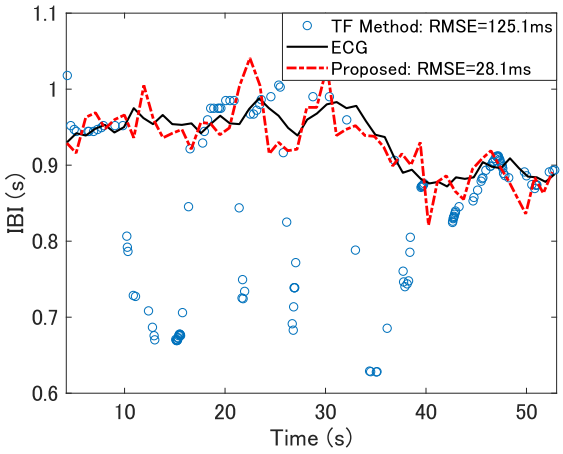


FIGURE 12. TF Method vs. the proposed method for estimating IBIs for P3.

TABLE 1. Experiment settings and comparison results for seven participants

	P1	P2	P3	P4	P5	P6	P7	Avg.
Participant age	63	33	25	26	25	23	24	31.3
Participant BMI	22.3	25.3	20.7	24.2	25.4	22.2	19.0	22.7
RMSE of TF Method (ms)	17.3	116.2	125.1	8.8	15.2	32.4	144.5	65.6
Mean±STD of TF Method (ms)	13.5 ± 11.2	75.5 ± 87.0	78.3 ± 98.2	6.8 ± 5.6	12.6 ± 9.2	23.4 ± 22.7	95.2 ± 117.5	43.6 ± 50.2
RMSE of our method (ms)	4.5	26.0	28.1	5.4	13.7	18.3	48.5	20.6
Mean±STD of our method (ms)	3.7 ± 2.7	19.2 ± 17.7	22.4 ± 17.1	4.0 ± 3.5	11.6 ± 7.3	14.9 ± 10.7	33.3 ± 35.6	15.6 ± 13.5

inaccurate, which makes the tracking of the heartbeat a difficult task. Moreover, the RMSE of the TF Method is relatively large and unstable. By contrast, the proposed method can accurately estimate the IBIs, even in some difficult cases. Table I summarizes the comparison results. The results in the figures and the table verify that the proposed method achieves an overall better performance than the TF Method, which demonstrates the efficiency of the CNN-based approach in heartbeat estimation.

V. DISCUSSION

A. WHY THE CNN CAN ESTIMATE HEART IBI USING RADAR SIGNALS

The proposed CNN-based algorithm is used to learn the relationship between the radar and ECG-related triangular signals for estimating the heart rate. Each cardiac muscle depolarization followed by repolarization during each cardiac cycle causes small electric field changes in the skin, which can be captured by the ECG electrodes. Almost simultaneously, the heartbeat also causes small body movements, which can be detected by UWB radar. Therefore, there should be a relationship between the radar and ECG signals if they are measured from the same person at the same time. This relationship is highly nonlinear and difficult to elucidate with-

out using a machine learning approach. Because CNNs can learn such nonlinear relationships effectively, the proposed CNN-based algorithm performs better than the conventional non-machine learning method.

According to [45], ECG is more accurate than PPG in monitoring the heart rate. Thus, we use ECG data to train the CNNs. In future, it will also be important to study the performance of the proposed algorithm using PPG. Additionally, the reproducibility of the heartbeat patterns in ECG data has been established [46], [47]. Based on this, we use the CNNs to learn the relationship between the radar and ECG data so that the radar signals can be associated with the heartbeat patterns for personal identification and heart rate estimation.

B. WHY THE PROPOSED METHOD CAN OVERCOME THE SYNCHRONIZATION DEVIATION

Fig. 13 shows an example with a non-zero synchronization deviation. The first row contains raw radar signals that were used as the input of the CNNs. The second row contains ECG signals that were measured as a reference of the ground truth. The third row contains two triangular waves transformed from the ECG signals shown in the second row. The fourth row contains the outputs of the CNNs. Fig. 13 shows that the output of the CNNs trained on radar signals shifted to the right slightly compared with the triangular waves transformed from the ECG, which demonstrates that the synchronization deviation occurred during the training of the CNNs of this subject. The dashed blue and purple vertical lines show this small deviation visually. Even though this deviation existed, the heartbeat could be estimated correctly. The deviation was consistent over time, and the IBI was calculated from the intervals between adjacent peaks or valleys. Therefore, the IBI was not influenced by the synchronization deviation. If we did not use five parallel CNNs, the system could not produce smooth waves when the synchronization of the radar and ECG signals deviated. The results demonstrate that our parallel structure is effective in estimating the heartbeat, even if the radar and ECG are not synchronized completely.

C. WHY IT IS UNNECESSARY TO REMOVE THE RESPIRATION COMPONENTS

In previous radar-based heartbeat monitoring studies, the respiration component interfered with the heartbeat component, which lowered the accuracy of the heart rate estimation. However, this analysis does not apply to the proposed CNN-based algorithm. In this section, we demonstrate how respiration affects the estimation of the heart rate when using the conventional and proposed CNN-based methods. For this purpose, we apply a high-pass filter (HPF) with a cut-off frequency of 0.27 Hz to radar signal $d(t)$ and obtain $d_{\mathrm{H}}(t) = \mathrm{HPF}_{0.27\mathrm{Hz}}[d(t)]$, which contains a high-frequency component. We then obtain $d_{\mathrm{L}}(t) = d(t) - d_{\mathrm{H}}(t)$, which contains a low-frequency component. Note that $d_{\mathrm{H}}(t)$ contains most of the heartbeat component, $d_{\mathrm{L}}(t)$ contains most of the respiration component, and $d(t)$ contains both components. Fig. 14 shows examples of $d(t)$, $d_{\mathrm{H}}(t)$, and $d_{\mathrm{L}}(t)$. Fig. 15 shows the heart IBIs estimated using $d(t)$, $d_{\mathrm{H}}(t)$, and $d_{\mathrm{L}}(t)$, respectively. These results show that the IBI estimated using $d(t)$ is more accurate than the IBIs estimated using $d_{\mathrm{L}}(t)$ or $d_{\mathrm{H}}(t)$, which implies that it is not necessary to remove the respiration component when using the proposed CNN-based algorithm, which provides new insights into the conventional belief of the negative effect of respiration components [40].

Using $d_{\mathrm{L}}(t)$ containing only the respiration component can rarely estimate the heart rate, which indicates that the heartbeat component is still more important than the respiration component, even for the proposed CNN-based algorithm. Table II summarizes the comparison of the accuracies of the

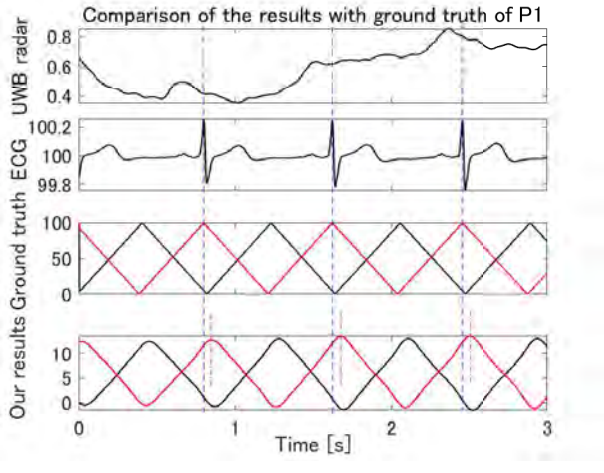


FIGURE 13. Comparison of the ground truth with the results from CNNs: The first and second rows show the radar and ECG, respectively. The red and black curves in the third row show two rectangular waves corresponding to the R and S waves, respectively, of the ECG, and these waves are the ground truth. The red and black curves in the fourth row show the outputs of the trained CNNs.

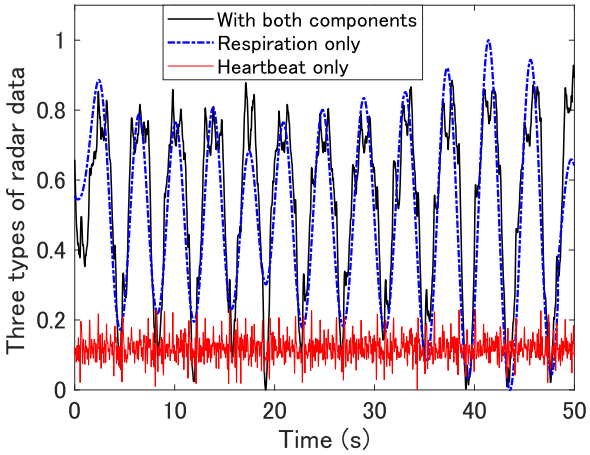


FIGURE 14. Three different types of radar data: (1) data that contain both respiration and heartbeat components; (2) data that contain the respiration component only; and (3) data that contain the heartbeat component only.

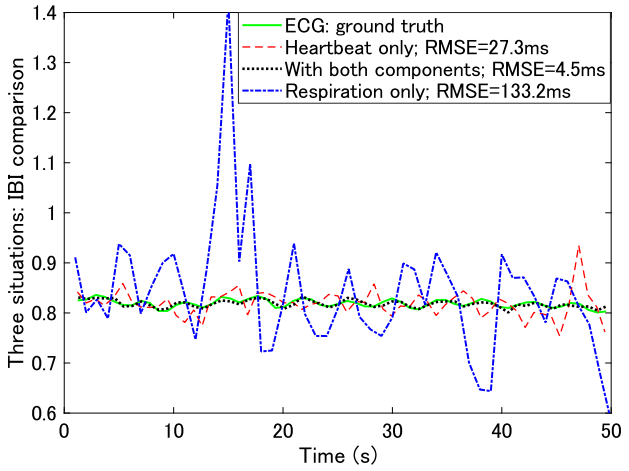


FIGURE 15. Comparison of results obtained from the three different types of data.

TABLE 2. RMSE comparisons for the results obtained by filters 1 and 2 (without and with the respiration component, respectively)

All RMSEs are in ms	P1	P2	P3	P4	P5	P6	P7
TF Method: without respiration	17.3	116.2	125.1	8.8	15.2	32.4	144.5
TF Method: with respiration	97.8	52.2	NA	9.2	24.6	67.0	54.1
Our method: without respiration	27.3	84.3	95.6	10.6	20.6	34.4	97.7
Our method: with respiration	4.5	26.0	28.1	5.4	13.7	18.3	48.5

proposed CNN-based algorithm and the TF Method using $d(t)$ and $d_{\mathrm{H}}(t)$, where $d_{\mathrm{L}}(t)$ is excluded from the comparison because neither algorithm can estimate the heart rate accurately using $d_{\mathrm{L}}(t)$. Note that the second row of P3 is not available (NA) because of the instability of the TF Method in this specific case. The table shows that the respiration component improved the accuracy of the TF Method in most cases, whereas removing the respiration component degraded the accuracy of the proposed algorithm in all cases.

These results can be explained by the mechanism of the TF Method; that is, the TF Method uses the feature points of a radar signal. When the radar signal contains both respiration and heartbeat components, the TF Method cannot distinguish the features of the heartbeat and respiration signals correctly, which results in its low accuracy. In cases P2 and P7 in Table II, removing the respiration component degraded the accuracy of the TF Method. This is because the high-pass filter erroneously excluded the heartbeat component in addition to the respiration component in these two cases. In general, it is not easy to determine the optimum cut-off frequency to separate the heartbeat and respiration components because these components often overlap in the frequency domain. By contrast, it is noteworthy that the respiration component may benefit the proposed CNN-based algorithm, which means that adjusting the cut-off frequency of the filter is not necessary when using the proposed CNN-based algorithm.

D. MOTION OF THE TARGET PERSON

As stated in Section III, we assumed that the proposed CNN-based algorithm was intended to be applied to a user in static scenarios, where the SNR remains low, and mostly unchanged. In general, radar-based heartbeat monitoring of an active user in motion is challenging, regardless of whether conventional methods or machine learning methods are used. Because the purpose of this study is to show for the first time the effectiveness of machine-learning-based heartbeat monitoring and personal identification using UWB radar, tackling this challenging problem is out of the scope of the present paper. In future studies, it will be important to investigate the performance of the proposed CNN-based algorithm when the target person is in motion.

VI. CONCLUSION

We proposed a CNN-based algorithm to accurately estimate the heartbeat from a UWB radar signal. The proposed method uses supervised learning with multiple CNNs trained using radar-ECG signal pairs as a training set. The experimental results demonstrate that the proposed method not only improves the accuracy in estimating the heartbeat, but also simplifies the entire signal processing because the proposed method does not need preprocessing such as multiple-parameter-based filtering. Interestingly, the proposed method was found to perform better when the data contain respiration components, which degrade the accuracy of conventional algorithms. The smallest, largest, and average RMS errors of the heart IBI estimated using the TF Method were 8.8, 144.5, and 65.6 ms, respectively. By contrast, the smallest, largest, and average RMS errors of the heart IBI estimated using the proposed CNN-based algorithm were 4.5, 48.5, and 20.6 ms, respectively. The degree of accuracy improvement was 4.5 times in the best case, 1.1 times in the worst case, and 3.2 times on average. Because the proposed method can identify the individual among several possible participants, the heart rate is measured and associated with the individual record using only radar signals without the need for any additional information and techniques.

A limitation of this research is that the proposed model was designed for static scenarios only. In our future work, we will research how to process different scenarios, for example, when the subject is in motion, at different orientations, and under stress. Additionally, we plan to visualize the data flows inside the CNN in the next stage of our research.

VII. ACKNOWLEDGMENT

The ethical aspects of the current work were approved by the Ethics Committee, Graduate School of Informatics, Kyoto University, under Registration No. 2016-011: Human vital monitoring using UWB radar. We thank Maxine Garcia, PhD, from Edanz Group (www.edanzediting.com/ac) for editing a draft of this manuscript.

REFERENCES

[1] L. Mesin, "Heartbeat monitoring from adaptively down-sampled electrocardiogram," *Comput. Biol. Med.*, vol. 84, no. 2017, pp. 217–225, 2017.

[2] M. Weenk et al., "Continuous monitoring of vital signs using wearable devices on the general ward: Pilot study," *JMIR Mhealth and Uhealth*, vol. 5, no. 7, pp. 1–12, 2017.

[3] E. Kaappa et al., "Performance analysis of novel flexible electrodes for wearable ECG/heart rate monitoring," in Proc. EMBEC-NBC, Tampere, Finland, 2017, pp. 237–240.

[4] S. Kim et al., "Heart rate detection during sleep using a flexible RF resonator and injection-locked PLL sensor," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 11, pp. 2568–2575, 2015.

[5] A. Temko, "Accurate heart rate monitoring during physical exercises using PPG," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 9, pp. 2016–2024, 2017.

[6] A. Martin and J. Voix, "In-ear audio wearable: Measurement of heart and breathing rates for health and safety monitoring," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 6, pp. 1256–1263, 2018.

[7] X. Hui and E. Kan, "Monitoring vital signs over multiplexed radio by near-field coherent sensing," *Nat. Electron.*, vol. 1, pp. 74–78, 2018.

[8] M. Villarroel et al., "Non-contact vital sign monitoring in the clinic," in Proc. IEEE FG, Washington, DC, USA, 2017, pp. 278–285.

[9] M. Kumar et al., "DistancePPG: Robust non-contact vital signs monitoring using a camera," *Biomed. Opt. Express*, vol. 6, no. 5, pp. 1565–1588, 2015.

[10] A. Qayyum et al., "Estimation of non-contact smartphone video-based vital sign monitoring using filtering and standard color conversion techniques," in Proc. IEEE LSC, Sydney, NSW, Australia, 2017, pp. 202–205.

[11] S. Sanyal and K. Nundy, "Algorithms for monitoring heart rate and respiratory rate from the video of a user's face," *IEEE J. Transl. Eng. Health Med.*, vol. 6, pp. 2168–2372, 2018.

[12] L. Chioukh et al., "Integrated radar systems for precision monitoring of heartbeat and respiratory status," in Proc. IEEE APMC, Singapore, 2009, pp. 405–408.

[13] G. Vinci et al., "Six-port radar sensor for remote respiration rate and heartbeat vital-sign monitoring," *IEEE Trans. Microw. Theory Tech.*, vol. 61, no. 5, pp. 2093–2100, 2013.

[14] J. Lai et al., "Wireless sensing of human respiratory parameters by low-power ultrawideband impulse radio radar," *IEEE Trans. Instrum. Meas.*, vol. 60, no. 3, pp. 928–938, 2011.

[15] H. Tan et al., "Non-contact heart rate tracking using Doppler radar," in Proc. ICSAI, Yantai, China, 2012, pp. 1711–1714.

[16] W. Hu et al., "Noncontact accurate measurement of cardiopulmonary activity using a compact quadrature Doppler radar sensor," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 3, pp. 725–735, 2014.

[17] A. Rahman et al., "Doppler radar techniques for accurate respiration characterization and subject identification," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 8, no. 2, pp. 350–359, 2018.

[18] T. Sakamoto et al., "Feature-based correlation and topological similarity for interbeat interval estimation using ultra-wideband radar," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 4, pp. 747–757, 2016.

[19] T. Sakamoto et al., "Remote heartbeat monitoring from human soles using 60-GHz ultra-wideband radar," *IEICE Electron. Express*, vol. 12, no. 21, Art.no. 20150786, 2015.

[20] T. Sakamoto et al., "Accurate heartbeat monitoring using ultra-wideband radar," *IEICE Electron. Express*, vol. 12, no. 3, Art.no. 20141197, 2015.

[21] I. Walterscheid et al., "Respiration and heartbeat monitoring using a distributed pulsed MIMO radar," in Proc. EMBC, Seogwipo, South Korea, 2017, pp. 3449–3452.

[22] H. S. Cho et al., "Detection of heart rate through a wall using UWB impulse radar," *J. Healthc Eng.*, vol. 2018, pp. 1–7, 2018.

[23] Y. Lee et al., "A novel non-contact heart rate monitoring using impulse-radio ultra-wideband (IR-UWB) radar technology," Sci. Rep., 2018. [Online]. Available: DOI:10.1038/s41598-018-31411-8

[24] L. Ren et al., "Noncontact multiple heartbeats detection and subject localization using UWB impulse Doppler radar," *IEEE Microw. Wirel. Compon. Lett.*, vol. 25, no. 10, pp. 690–692, 2015.

[27] W. Hung et al., "Real-time and noncontact impulse radio radar system for μm movement accuracy and vital-sign monitoring applications," *IEEE Sen. J.*, vol. 17, no. 8, pp. 2349–2358, 2017.

[25] K. Naishadham et al., "Estimation of cardiopulmonary parameters from ultra wideband radar measurements using the state space method," *IEEE Trans. Biomed. Circuits Syst.*, vol. 10, no. 6, pp. 1037–1046, 2016.

[26] L. Ren et al., "Phase-based methods for heart rate detection using UWB impulse Doppler radar," *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 10, pp. 3319–3331, 2016.

[28] S. Nahar et al., "An electromagnetic model of human vital signs detection and its experimental validation," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 8, no. 2, pp. 338–349, 2018.

[29] E. Schires et al., "Vital sign monitoring through the back using an UWB impulse radar with body coupled antennas," *IEEE Trans. Biomed. Circuits Syst.*, vol. 12, no. 2, pp. 292–302, 2018.

[30] K. Shyu et al., "Detection of breathing and heart rates in UWB radar sensor data using FVPIEF based two-layer EEMD," *IEEE Sens. J.*, vol. 19, no. 2, pp. 774–784, 2018.

[31] U. R. Acharya et al., "A deep convolutional neural network model to classify heartbeats," *Comput. Biol. Med.*, vol. 89, pp. 389–396, 2017.

[32] U. R. Acharya et al., "Deep convolutional neural network for the automated detection and diagnosis of seizure using ECG signals," *Comput. Biol. Med.*, vol. 100, pp. 270–278, 2018.

[33] P. Rajpurkar et al., "Cardiologist-level arrhythmia detection with convolutional neural networks," *Nature Med.*, vol. 25, pp. 65–69, 2019.

[34] S. Kiranyaz et al., "Real-time patient-specific ECG classification by 1-D convolutional neural networks," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 3, pp. 664–675, 2016.

[35] O. Yildirim, "A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification," *Comput. Biol. Med.*, vol. 96, pp. 189–202, 2018.

[36] W. Yin et al., "ECG monitoring system integrated with IR-UWB radar based on CNN," *IEEE Access*, vol. 4, pp. 6344–6351, 2016.

[37] D. Guermandi et al., "A 79GHz 2 $\times$ 2 MIMO PMCW radar SoC in 28nm CMOS," *IEEE J. Solid-State Circuits*, vol. 52, no. 10, pp. 2613–2626, 2017.

[38] T. Sakamoto et al., "Measurement of instantaneous heart rate using radar echoes from the human head," *Electron. Lett.*, vol. 54, no. 14, pp. 864–866, 2018.

[39] J. T. Springenberg et al., "Striving for simplicity: The all convolutional net," in ICLR (Workshop Track), 2015.

[40] I. V. Mikhelson et al., "Remote sensing of heart rate and patterns of respiration on a stationary subject using 94-GHz millimeter-wave interferometry," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 6, pp. 1671–1677, 2011.

[41] T. Sakamoto et al., "Noncontact measurement of the instantaneous heart rate in a multi-person scenario using X-band array radar and adaptive array processing," *IEEE Trans. Emerg. Sel. Topics Circuits Syst.*, vol. 8, no. 2, pp. 280–293, 2018.

[42] Unwrap function in MATLAB. [Online]. Available: http://cens.ioc.ee/local/man/matlab/techdoc/ref/unwrap.html

[43] MZ. Alom et al., "The history began from AlexNet: A comprehensive survey on deep learning approaches," arXiv preprint, 2018. [Online]. Available: https://arxiv.org/abs/1803.01164

[44] S. V. Georgakopoulos et al., "Convolutional neural networks for toxic comment classification," in Proc. SETN, Patras, Greece, 2018, pp. 1–6.

[45] "ECG vs PPG for Heart Rate Monitoring: Which is Best?," 2015. [Online]. Available: http://neurosky.com/2015/01/ecg-vs-ppg-for-heart-rate-monitoring-which-is-best

[46] W. Louis et al., "Continuous authentication using one-dimensional multi-resolution local binary patterns (1DMRLBP) in ECG biometrics," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 12, pp. 2818–2832, 2016.

[47] I. Odinaka et al., "ECG biometric recognition: a comparative analysis," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 6, pp. 1812–1824, 2012.