ULTRA WIDE-BAND RADAR FOR THE REAL-TIME MONITORING OF HEART RATE USING CONVOLUTIONAL NEURAL NETWORK

ỨNG DỤNG MẠNG NƠ RON TÍCH CHẬP CHO HỆ THỐNG THEO DÕI NHỊP TIM THEO THỜI GIAN THỰC SỬ DỤNG RA ĐA BĂNG THÔNG SIÊU RỘNG

LE DANG KHANH
Faculty of Engineering, Vietnam Maritime University
Email: ledangkhanh@vimaru.edu.vn

Abstract

Ultra-wideband (UWB) radars are getting much attention for maritime applications of smart and luxury ships in which UWB radar could be integrated into Bridge Navigational Watch & Alarm System - BNWAS. One of the interesting applications of UWB radar is vital signs measurement, which is a contactless method. UWB radar measures respiration and heartbeat rate by the motion of thorax for detecting and checking the state of people on the bridge. However, the motion of the thorax caused by the heartbeat is usually low intensity and easily gets noisy and perturbed by a non-stationary signal. Due to this, an architecture built by a convolutional neural network is developed and modified to monitor heart rate using a contactless ultra wide-band (UWB) radar. The preprocessing part including many steps is necessary to clean raw signals from UWB radar. In this study, the evaluation metrics included a root mean square error of 11.34, a mean absolute error of 8.98, a standard deviation of the estimated signal of 4.05, and a percentage error of average HR at 5.77%. The proposed model could capture HR and is expected to be used for monitoring health and psychological status.

Keywords: *Vital sign, UWB radar, heart-rate monitoring, convolutional neural network, Real-time monitoring.*

Tóm tắt

Các radar băng thông siêu rộng (UWB) đang nhận được nhiều sự quan tâm đối với các ứng dụng hàng hải trên các tàu thông minh và tàu hạng sang, trong đó radar UWB có thể được tích hợp vào Hệ thống cảnh báo và giám sát cầu dẫn đường - BNWAS. Một trong những ứng dụng thú vị của radar UWB là đo các tín hiệu sống, đây là phương pháp không tiếp xúc. Radar UWB đo nhịp thở và nhịp tim bằng chuyển động của lồng ngực để phát hiện và kiểm tra trạng thái của người trong ca trực. Tuy nhiên, chuyển động của lồng ngực do nhịp tim gây ra thường có cường độ thấp và dễ bị nhiễu và nhiễu loạn bởi tín hiệu không cố định. Do đó, một thuật toán được phát triển dựa trên mạng nơ ron tích chập để theo dõi nhịp tim không tiếp xúc bằng radar băng thông siêu rộng (UWB). Phần tiền xử lý bao gồm nhiều bước cần thiết để xử lý tín hiệu thô từ radar UWB. Trong nghiên cứu này, các số liệu đánh giá bao gồm sai số bình phương trung bình gốc là 11,34, sai số tuyệt đối trung bình là 8,98, độ lệch chuẩn của tín hiệu ước tính là 4,05 và sai số phần trăm của HR trung bình là 5,77%. Mô hình đề xuất có thể đo được nhịp tim và phục vụ cho công tác theo dõi tình trạng sức khoẻ, tâm lý của người được theo dõi.

Từ khóa: *Tín hiệu sống, radar UWB, Giám sát nhịp tim, Mạng nơ ron tích chập, Giám sát thời gian thực.*

Abbreviations

HR	Heart-rate
UWB	Ultra wide-band
CNN	Convolutional neural network
RMSE	Root mean square error
MAE	Mean absolute error
PPG	Photoplethysmography

1. Introduction

The UWB radar has recently been used in indoor applications in smart homes, or smart cities due to its advantages. A UWB radar has low power consumption, and simple architecture, but provides rich information about the spatial environment and high resolution. Therefore, it is widely used in indoor applications such as driver safety assistant [1], people counting [2, 3], through wall human detection [4]. Besides, UWB radar is very sensitive in that it can measure the tiny motion from thorax in the breathing and cardiac activities, so it can monitor HR.

TẠP CHÍ ISSN: 1859-316X
KHOA HỌC CÔNG NGHỆ HÀNG HẢI
JOURNAL OF MARINE SCIENCE AND TECHNOLOGY

Many methods have been published to track the HR. For example, a CNN model was used to monitor vital signs, including HR and respiration rate, based on impulse radio ultrawide-band radar during sleep [5]. The paper provides a method of using both radar signals and applying a continuous wavelet transform to have information about the time domain to monitor HR for a long period. The model includes two parts, one is 1D CNN to learn the radar signal and the other is 2D CNN to capture information of continuous wavelet transform signal. The average MAE for respiration rate and HR are 2.67 and 4.78. Although the method could monitor both respiration rate and HR, its result in monitoring HR has sections in which the MAE was larger than the HR estimated.

In this work, a much simpler architecture of the CNN model was developed to estimate HR. In this study, we aim to get more information by taking several consecutive signals concerning fast time as the input instead of only one signal with the richest information. The preprocessing part included band-pass filtering, frame stacking, clutter removal, person detection, and min-max normalization to have a better dataset for the model. The CNN architecture used a 1D convolutional layer as the core.

2. Methods

2.1. Experimental set-up

A UWB radar in front of a person's heart, and a PPG, which was treated as a ground truth signal, were used to collect vital signs simultaneously, as shown in Figure 1.

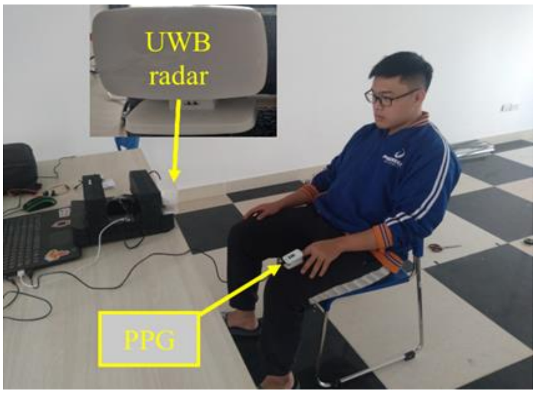

Figure 1. Experimental set-up

In this paper, we use UWB radar produced by UMAIN company with a center frequency of 4.6 GHz and bandwidth 500 MHz, with the setting of sampling rate F_S of 44.5 frame/second. The UWB radar is placed on the table which is 1m away from the person.

2.2. Data collection and preprocessing

There were 18 people taking part in collecting data with both male and female participants. To have a reliable result, consider the data of 14 people as a training dataset, and the rest as the valid dataset. After this, the raw dataset was created, and were able to move on to the preprocessing part, as shown in Figure 2. To preprocess the raw signal, firstly, apply the band-pass filter, with a range of frequency between 0.01Hz and 5Hz, which can capture heart rate as the normal frequency of heart rate is equivalent to 0.6Hz to 2Hz. Then, to provide more information and increase stability, the dataset is stacked in a window size of 128 signals, with a sliding step of 11 frames. Next is to remove clutter by applying a median filter and detecting the distance of a person in front of the UWB radar by finding the index of the highest standard deviation between data points concerning fast time. After that, the input radar signal of the model is defined as the 1D signal, with a window size, of 128, times the range around the highest standard deviation index, in this paper, it is set at 30. So, each radar sample has a size of 128×30 corresponding to 6.4s × 600mm. And the last step, to increase the training speed and the stability of the model and keep the signal shape, apply min-max normalization. About the label dataset, the signal received from PPG in the time domain is converted into beat per minute domain by applying Chirp Z-transform [5].

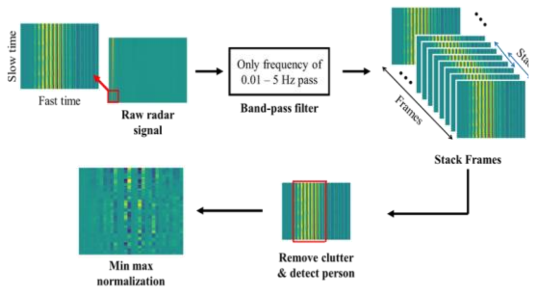

Figure 2. Preprocessing steps

2.3. CNN model

A Convolutional Neural Network (CNN) represents a specialized type of deep learning architecture commonly employed in the field of Computer Vision. Computer vision, a subdomain of Artificial Intelligence, empowers computers to comprehend and interpret visual data, including images.

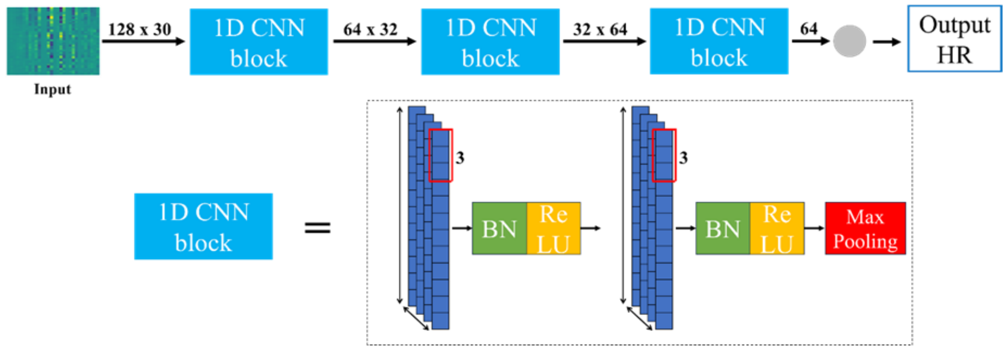

Figure 3. CNN model. Where the 1D CNN layer uses a kernel size of 3, BN is batch normalization, ReLU is the activation layer; the last 1D CNN block uses Average Pooling instead of Max Pooling as shown

In the realm of Machine Learning, Artificial Neural Networks exhibit impressive performance. These networks find application across diverse datasets, encompassing images, audio, and text. Depending on the specific task, different types of neural networks are employed. For instance, when predicting word sequences, Recurrent Neural Networks (RNNs)-particularly Long Short-Term Memory (LSTM) networks-are commonly used. Similarly, for image classification tasks, Convolutional Neural Networks (CNNs) are the go-to choice.

In a regular Neural Network there are three types of layers:

- Input Layers: These layers receive input data for our model. The number of neurons in this layer corresponds to the total number of features in our dataset (e.g., the number of pixels in an image).

- Hidden Layers: The input from the Input layer is subsequently fed into the hidden layers. Depending on our model and data size, there can be multiple hidden layers. Each hidden layer may contain varying numbers of neurons, typically exceeding the number of input features. The output from each hidden layer is computed through matrix multiplication with learnable weights specific to that layer. Additionally, learnable biases are added, followed by an activation function. This nonlinearity introduced by the activation function is crucial for the network's expressive power.

- Output Layer: The output from the hidden layers is then directed to a logistic function, such as the sigmoid or softmax function. These functions convert the raw output values for each class into probability scores, facilitating classification decisions.

In this research, we will investigate into constructing a fundamental building block for CNNs. There are many changes in the number of CNN blocks, filter size, and segment window size, and the best performance was applied. Figure 3 depicts the general block diagram of the network. There are three CNN blocks, each block has a sequence of 2 convolutional layers (painted blue), 2 batch normalize layers, 2 ReLU activation layers, max pooling layer, except for the last block, which uses the average pooling layer instead. The 2 convolutional layers use the same number of filters, and after each block, that number doubles, from 16 to 64, consequently. The reason for that is to let the convolutional layer learn the features as much as possible before down sampling the dataset; and as moving forward in CNN blocks, the layer needs to capture more complex patterns, and as many patterns as possible. Each convolutional layer contains the regularization to help prevent overfitting. A batch-normalized layer is used in each block to increase the training speed and the stability of the model.

2.4. Results

Table 1 shows the evaluation metrics of models, including Root Mean Square Error (RMSE), and Mean Absolute Error (MAE), the gap between the ground truth signal and the estimated signal of HR from the valid dataset is shown in percentage and standard deviation (Std).

Evaluation metric formulas:

RMSE:

$$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (1)$$

TẠP CHÍ KHOA HỌC CÔNG NGHỆ HÀNG HẢI ISSN: 1859-316X
JOURNAL OF MARINE SCIENCE AND TECHNOLOGY

MAE:

$$\frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (2)$$

Std:

$$\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \frac{1}{n}\sum_{j=1}^{n}x_j\right)^2} \quad (3)$$

Percentage:

$$\frac{\sum_{i=1}^{n}|y_i - \hat{y}_i|}{\sum_{j=1}^{n}y_j} \quad (4)$$

where y *is ground truth signal;* $\hat{y}$ *is the estimated signal; and* $x = |y - \hat{y}|$.

Compared to the proposed model, which has 3 CNN blocks, if the number of CNN block is increased to 4, all the evaluation values are higher, which depicts worse result. On the other hand, lower the number of filters or the number of convolutional layer of each CNN block may result some better values of evaluation metrics (values of RMSE, MAE of the model near the last row are 10.37 and 8.14, compared to 11.34 and 8.98 of proposed model). However, due to the lack of parameters, the model has less ability to track HR, which results less reliable prediction. Generally, increasing the number of blocks or the starting number of filters will increase the error or overfitting where decreasing them can cause underfitting.

Figure 4 depicts the ground truth and estimated HR. Basically, the HR prediction does follow the general trend of the ground truth signal. There are various points that depict the gap between the estimated signal and the ground truth signal. This could be due to the noise when using PPG (user does move his hand for a while), and the limitations model's ability to track the sudden change in HR.

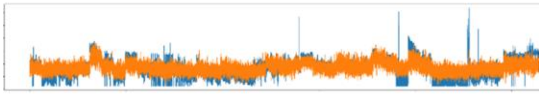

Figure 4. The result for valid signal

3. Conclusion

Heart rate (HR), a fundamental physiological parameter, plays a crucial role in assessing overall health and monitoring disease progression. In pursuit of non-invasive and efficient heart rate monitoring, an innovative architecture leveraging a convolutional neural network (CNN) has been meticulously designed and adapted for use with a contactless ultra-wideband (UWB) radar. The UWB radar captures raw signals, necessitating a comprehensive preprocessing pipeline to enhance signal quality.

In this study, we focus on extracting valuable insights by analyzing consecutive signals within a short time frame. These sequential signals serve as input to the CNN model, which robustly estimates heart rate. The evaluation of our model reveals promising performance metrics:

- RMSE: Achieving an RMSE of 11.34, our model demonstrates accurate estimation.

- MAE: With an MAE of 8.98, our approach minimizes deviations from ground truth.

- Std: The estimated signal exhibits stability, with a standard deviation of 4.05.

- Percentage Error of Average HR: Our model

Table 1. Evaluation metrics

	RMSE	MAE	Std	Percentage
Proposed model	11.34	8.98	4.05	5.77%
4 blocks, 16 filters start	13.98	11.51	8.87	5.78%
4 blocks, 1 conv layer each block, 16 filters start	11.90	9.59	7.13	5.98%
3 blocks, 32 filters start	14.59	11.93	7.17	11.07%
3 blocks, 1 conv layer each block, 16 filters start	10.37	8.14	4.70	5.99%
3 blocks, 1 conv layer each block, 32 filters start	11.84	9.46	6.22	4.53%

where 1 conv layer each block, only one combination of conv layer, BN, ReLU is used; filter start, the starting number of filter; RMSE, root mean square error; MAE, mean absolute error; std, standard deviation of estimated signal; percentage, error between average HR value of ground truth signal and estimated signal.

maintains precision, with an average heart rate percentage error of 5.77%.

The proposed CNN-based framework effectively captures heart rate dynamics and holds promise for practical applications in health monitoring. Its potential deployment in clinical settings could revolutionize disease management and enhance patient care.

Acknowledgment

This research is funded by Vietnam Maritime University under grant number: **DT23-24.16**.

REFERENCES

[1] Zhao, P, Lu, CX, Wang, J, Chen, C, Wang, W, Trigoni, N & Markham (2021), *Human tracking and identification through a millimeter wave radar,* Ad Hoc Networks, Vol.116, 102475.

[2] Kyou-Kai Shyu, Luan-Jiau Chiu, Po-Lei Lee, Tzu-Han Tung (2018), *Detection of Breathing and Heart Rates in UWB Radar Sensor Data using FVPIEF Based Two-Layer EEMD*, IEEE Sensors Journal PP (99):1-1.

[3] Y. Zhang, X. Li, R. Qi, Z. Qi, and H. Zhu (2020), *Harmonic Multiple Loop Detection (HMLD) Algorithm for Not-Contact Vital Sign Monitoring Based on Ultra-Wideband (UWB) Radar*, in IEEE Access, Vol. 8, pp.38786-38793.

[4] Choi, S.H.; Yoon, H (2023). *Convolutional Neural Networks for the Real-Time Monitoring of Vital Signs Based on Impulse Radio Ultrawide-Band Radar during Sleep*. Sensors, Vol. 23 (6), 3116. https://doi.org/10.3390/s23063116.

[5] M. Le, V. S. Luong, K. Dang Nguyen, T. D. Le and D. -K. Le (2023), *Multivariate Signal Decomposition for Vital Signal Extraction using UWB Impulse Radar*, 2023 IEEE Statistical Signal Processing Workshop (SSP), Hanoi, Vietnam, pp.290-294.
doi: 10.1109/SSP53291.2023.10208009.

Received:	03 April 2024
Revised:	12 April 2024
Accepted:	16 April 2024