

# Efficient J Peak Detection From Ballistocardiogram Using Lightweight Convolutional Neural Network\*

Yongfeng Huang<sup>1</sup>, Tianchen Jin<sup>1</sup>, Chenxi Sun<sup>1</sup>, Xueyang Li<sup>1</sup>, Shuchen Yang<sup>2</sup>, Zhiming Zhang<sup>2</sup>

**Abstract**—Ballistocardiogram (BCG) is a non-contact and non-invasive technique to obtain physiological information with the potential to monitor Cardio Vascular Disease (CVD) at home. Accurate detection of J-peak is the key to get critical indicators from BCG signals. With the development of deep learning methods, many researches have applied convolution neural network (CNN) and recurrent neural network (RNN) based models in J-peak detection. However, these deep learning methods have limitations in inference speed and model complexity. To improve the computational efficiency and memory utilization, we propose a robust lightweight neural network model, called JwaveNet. Moreover, in the preprocessing stage, J-peaks are re-modeled by a new transformation method based on their physiological meaning, which has been proven to increase performance. In our experiment, BCG signals, including four different sleeping positions, were collected from 24 subjects with synchronous electrocardiogram (ECG) signals. The experiment results have shown that our lightweight model greatly reduces latency and model size compared to other baseline models with high detecting accuracy.

## I. INTRODUCTION

Ballistocardiogram (BCG) is a measure of repetitive movements of the human body caused by heartbeat and acceleration of blood ejected in vessels [1]. Attributed to the mechanism of BCG, abundant physiological information of the human body can be obtained from BCG through the non-invasive and unobtrusive way, which can greatly reduce the psycho-physiological discomfort of patients.

“I”, “J”, and “K” waves are typical of BCG recordings and have been proved with clinical significance in cardiovascular parameters [2]. Especially, the J-wave is the largest headward wave that occurs late in systole. And the most prominent peak in each J-wave is named J-peak, which always represents the heartbeat as R-peak in ECG. Meanwhile, the location on J-peaks is also the key to calculate pulse transit time (PTT) [3]. Therefore, the precise localization of the J-peak is fundamental to obtain Inter-beat interval (IBI) estimation, heart rate (HR) estimation, cuff-less blood pressure (BP) monitoring and other physiological information.

In recent years, a large number of deep learning based-algorithms have been applied in the field of annotation of specific waves in biomedical signals [4-8]. Most typically, CNN and RNN methods are widely used to localize the waves both in ECG and BCG. Some studies have used a U-Net-like network to annotate ECG waves [4]. Similarly,

IJK segments in the BCG signal were detected using U-Net, and the J peaks were found in the maximum of the IJK probability sequence [5]. To improve the accuracy of peak detection in noisy data, RPnet, which was adapted from IncRes-Unet, showed good performance after Distance Transform (DT) in the different datasets [6]. Additionally, other studies intended to use RNN-based models for capturing temporal dependencies of ECG or BCG signal and achieved high accuracy in wave detection tasks [7,8].

However, most deep learning methods aim to improve the accuracy of peak detection without considering memory and computational power requirements. Moreover, among BCG peak detecting tasks using deep learning methods, few studies have been conducted to verify the effectiveness of the network models under new individuals. Therefore, it is uncertain if J-peaks can be detected on new subjects’ BCG signals. Inspired by Mini-Inception-Residual-Dense (MIRD) Net [9] and Mobilenets [10,11], we proposed a state-of-the-art lightweight neural network, called JwaveNet, which improves computational efficiency and saves storage space while preserving accuracy. In particular, to further enhance the detecting performance of the lightweight model in varieties of BCG waves, we build a J-wave model centered on every J-peak in the preprocessing stage. Finally, J-peaks will be detected utilizing simple post-processing. The main contributions are summarized as follows:

- We propose a robust lightweight network model for detecting J-peak of BCG efficiently and accurately;
- In the process of data preprocessing, a transformation method based on the physiological meaning of J-wave is applied to improve the performance of our lightweight neural network model.

## II. METHODOLOGY

### A. Data Collection

In this experiment, a polyvinylidene fluoride (PVDF) piezoelectric bed sensor consisted of 18 channels was used to acquire BCG signals with a sampling rate of 50 Hz. The experimental procedures were approved by the Institutional Review Board. And all volunteers provided informed consent for participation. At the time of measurement, the sensor was placed under the mattress, close to the subject’s head. And the detailed distribution information of the BCG signals on the 18 channels can be found in [12]. Meanwhile, the experiment was synchronized by the acquisition of ECG signals with a sampling rate of 500 Hz, representing the ground truth of J-peak detection. To simulate a real application

\*Research supported by Shanghai Yueyang Medtech Co., Ltd.

<sup>1</sup>Yongfeng Huang, Tianchen Jin, Chenxi Sun and Xueyang Li are with Computer Science and Technology Institute, Donghua University, Shanghai, 201620 China yfhuang@dhu.edu.cn

<sup>2</sup>Shuchen Yang and Zhiming Zhang are with Shanghai Yueyang Medtech Co., Ltd., Shanghai, 201203, China.

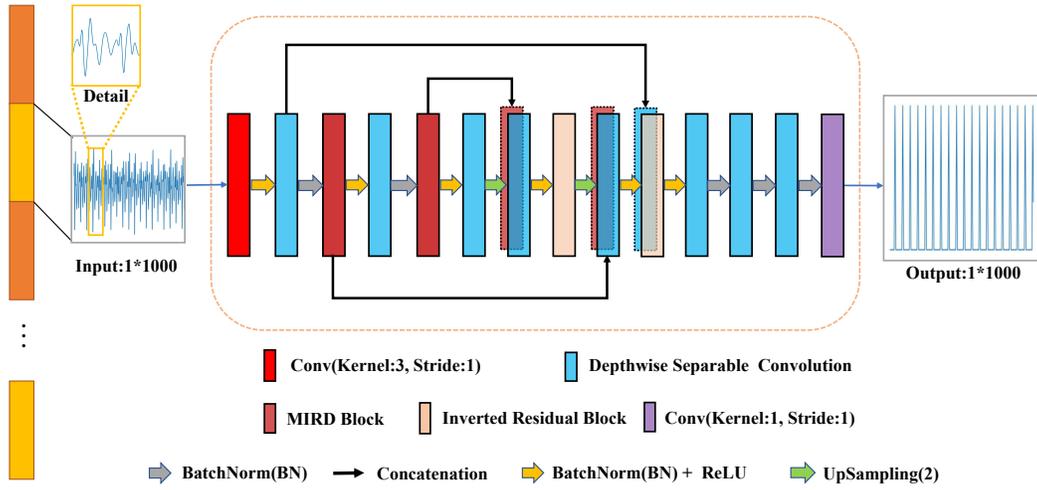


Fig. 1. Model Architecture.

scenario, we collected BCG data from 24 subjects in four sleeping positions (supine, left side-lying, right side-lying, and unrestricted position), and it took thirty minutes on each position with their most comfortable state. Ultimately, a total of approximately 48 hours of BCG signals with synchronous ECG signals were collected.

### B. Preprocessing

BCG signals acquired by the bed sensor were processed in the following steps. First, raw BCG signals were filtered by a seven-order-Butterworth filter with 2-10Hz 3dB cutoff frequency due to the fact that frequency of respiration component is generally lower than 2 Hz; Second, we selected one in eighteen channels of BCG signal which is approximately Gaussian as the object of J-peak detection, for BCG with fewer motion artifacts is more Gaussian distribution [13]. Third, 48 hours BCG signals and synchronous ECG signals were divided into 20-second segments, and we labeled J-peaks in BCG with an improved template matching algorithm [12], while R-peaks in ECG were labeled using the Pan-Tompkins method [14]. Then, J-peaks were corrected by R-peaks and checked by two experts. Finally, the Z-score was adopted to normalize BCG segments, the resulting BCG data were defined as  $X_k = \{x_1, x_2, x_3, \dots, x_t\}$ , and J-peaks were marked as  $Y_k = \{L_{Jpeak1}, L_{Jpeak2}, L_{Jpeak3}, \dots, L_{Jpeak_i}\}$ , where  $X_k, Y_k$  are the  $k^{th}$  BCG segment and its labels, respectively. And  $x_t$  represents the amplitude of normalized BCG at the  $t^{th}$ , while  $L_{Jpeak_i}$  is the location of the J peak at the  $i^{th}$ .

Considering that BCG waves vary greatly due to individual differences, we proposed a transformation method based on BCG physiological meaning to enhance the J-peak-detecting ability of the network model. According to analyze the genesis of BCG waves [2] and the statistics of the width of the IJK wave in every subject, it could be concluded that in the state of daily sleep, the width of the IJK wave (I-K interval) in each healthy individual is almost stable, so we centered on each J-peak label and roughly restore the J-wave to form a new label. In this task, we only set the I-J and J-K intervals to equal lengths. The specific transformation

algorithm of the J-peak label in each heartbeat interval is shown below:

$$y_t = \begin{cases} \frac{x_t - (L_{Jpeak_i} - fT/2)}{fT/2}, & \text{if } L_{Jpeak_i} - fT/2 < x_t \leq L_{Jpeak_i} \\ -\frac{x_t - (L_{Jpeak_i} + fT/2)}{fT/2}, & \text{if } L_{Jpeak_i} < x_t \leq L_{Jpeak_i} + fT/2 \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

Where  $f$  is the sampling rate of BCG signal,  $T$  denotes I-K interval. After the operation, the J-peak label is converted to  $Y'_k = \{y_1, y_2, y_3, \dots, y_t\}$ . Finally, the task of the neural network model is transformed from classifying J-peaks to predict rebuilding J-waves.

## III. MODEL ARCHITECTURE

To maintain the accuracy of new wave prediction with fewer parameters, we use a shallow symmetric architecture similar to MIRD-Net, which is composed of downsampling, upsampling and skip connection. In addition, we cancel the pooling operation in the downsampling process to alleviate the loss of information. Based on MIRD-Net, we propose a lightweight architecture called JwaveNet, which further increases the computational efficiency. The details of the architecture are shown in Fig. 1. The main blocks including Depthwise Separable Convolutions [10], MIRD blocks [9], and Inverted Residual Blocks [11] will be introduced below.

### A. Depthwise Separable Convolution

Depthwise Separable Convolution was proposed in MobilenetV1 [10], which decomposes a standard convolution operation into depthwise and pointwise convolutions, where the depthwise convolution operation filters each feature channel and the pointwise convolution regroups the features obtained from each channel. What's more, suggested by MobilenetV2 [11], we remove the non-linear ReLU activation function after pointwise convolution to reduce the loss of representational features caused by non-linear operations in low-dimensional space, and the Depthwise Separable Convolution is shown in Fig.2(a).

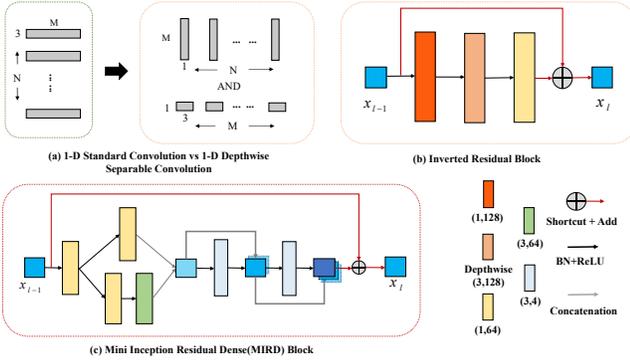


Fig. 2. The detail of Depthwise Separable Convolution (a), Inverted Residual Block (b) and Mini Inception Residual Dense (MIRD) Block (c). Note that the parameters in (1,128), (3,128), (1,64), (3,64) and (3,4) represent the kernel size and the number of output channels, respectively.

### B. Inverted Residual Block

The residual structure can fuse the feature information of the previous layer with the feature information of the current layer, which effectively alleviates the problem of gradient disappearance and overfitting. Different from the residual structure, in the inverted residual block illustrated in Fig.2(b), the depthwise separable convolution is used to extract the high-dimensional information more efficiently, and features in narrow layers are kept in a linear way.

### C. Mini Inception Residual Block

MIRD Block is a high-performance block shown in Fig.2(c), embedded Residual Block and Dense Block in inception architecture, which can alleviate the performance degradation caused by reduced network layers. Especially, inception architecture enables the network to acquire the features of BCG waves at different levels and enlarge the Receptive Field of the network. Moreover, Dense Block connects the current network layer to all previous layers, multiplying features of BCG waves that the network learned in the previous layer. Finally, the output is obtained by Residual operation. The number of feature channels in the MIRD Block is set to 4 or 64 to minimize parameters, and the kernel size of convolutions is 1 or 3 followed by Batch Normalization, ReLU activation.

### D. Post-processing

The lightweight network finally outputs a re-modeled J wave with a length of 20 seconds (same as the input length of BCG). Besides, to locate the J peaks on the J waves, we find all local maxima by a simple comparison of neighboring values with a minimal distance criterion.

## IV. EXPERIMENTS AND RESULTS

### A. Implementation Details

After removing obvious-motion-artifacts data that heartbeats could not be recognized by synchronous ECG, we got 6179 segments from 24 subjects. In order to prove that our network can achieve high accuracy in the J-peak detecting task of new subjects, we chose 1000 segments from randomly

4 subjects for training with 5-fold cross-validation. And other 20 subjects' 5179 segments were used as a test set.

In all experiments, we used the MSE loss function and Adam optimizer [14] with the batch size 10 over 100 epochs. The learning rate in training is 0.001 and the decay rate of learning rate is 0.1 after 20 epochs. All models were trained on a computer with Intel(R) Core (TM) i7-7700 CPU @ 3.60 GHz, Nvidia GeForce GTX 1080 Ti, 16 GB RAM, and Samsung SSD 850 EVO 500GB. All experiments were run under the Pytorch 1.6.0 version.

### B. Evaluation Metrics

The performance of J-peak detecting is evaluated by the precision, recall and F1 score, as shown below:

$$\begin{aligned} recall &= \frac{TP}{(TP + FN)}, \\ precision &= \frac{TP}{(TP + FP)}, \\ F1 - score &= \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \end{aligned} \quad (2)$$

where TP represents ground truth with a tolerance of 75ms, FP is the J-peaks detected out of the tolerant area of true peaks, and FN is the condition of failure to detect J-peaks during the permitted neighborhood.

In addition, mean absolute error (MAE) was used to evaluate the HR and IBI calculated through detected J-peaks. To measure the inference speed, running latency in the test set was calculated on both the central processing unit (CPU) and graphics processing unit (GPU). And the number of parameters was used to evaluate the model size.

### C. Result

We compared our JwaveNet with RPNNet [6], U-Net [15], shallower U-Net with 8 convolutional layers (U-Net-8), and MIRD-Net [9]. Additionally, we contrast the DT method used in [6] during preprocessing stage (Ours-DT) with the J-wave-transformation method proposed by us. As shown in Table I, JwaveNet with the new transformation method achieved a high F1-score of 0.9719, and MAE of IBI and MAE of HR are  $9.32 \pm 1.89$ ms and  $0.71 \pm 0.16$ bpm, respectively. Meanwhile, it is obvious that JwaveNet with the new transformation method has 3.6% higher F1-score while reducing MAE of IBI and HR by 22.81ms and 2.85bpm compared to the DT method (Ours-DT). This greatly proved our new transformation method is more suitable in the J-peak detection task.

Besides, Table II shows our model has a minimum number of parameters and high inference speed among the five models. Compared with RPNNet and U-Net, our network model, with no discernible difference in accuracy, has the parameters of only 3 / 1000 of IRD-Net and 3 / 200 of U-Net. In terms of the inference time, it took about 1.09s to process 5179 test segments (124597 beats) on GPU, while it took 62.32s on CPU. Namely, our model took only about 0.5ms to process a beat on CPU, which is approximately 184% faster than RPNNet, 52% faster than U-net and 16% faster than MIRD-Net with comparable accuracy.

TABLE I. Comparison of performance using different methods.

| Methods     | Recall        | Precision     | F1-score      | IBI <sub>MAE</sub><br>(ms) | HR <sub>MAE</sub><br>(bpm) |
|-------------|---------------|---------------|---------------|----------------------------|----------------------------|
| RPNet       | 0.9835        | 0.9876        | 0.9855        | 7.30±1.36                  | 0.56±0.08                  |
| U-Net       | 0.9755        | 0.9819        | 0.9787        | 8.48±1.07                  | 0.70±0.05                  |
| U-Net-8     | 0.9431        | 0.8955        | 0.9186        | 89.32±12.43                | 7.68±3.36                  |
| MIRD-Net    | 0.9545        | 0.9702        | 0.9623        | 9.80±3.56                  | 0.77±0.33                  |
| Ours-DT     | 0.9342        | 0.9375        | 0.9358        | 32.13±8.58                 | 3.56±1.82                  |
| <b>Ours</b> | <b>0.9657</b> | <b>0.9782</b> | <b>0.9719</b> | <b>9.32±1.89</b>           | <b>0.71±0.16</b>           |

TABLE II. Comparison of parameters and latency in different methods.

| Methods                     | RPNet  | U-Net | U-Net-8 | MIRD-Net | <b>Ours</b>  |
|-----------------------------|--------|-------|---------|----------|--------------|
| Parameters( $\times 10^6$ ) | 52.15  | 11.13 | 0.37    | 0.26     | <b>0.17</b>  |
| Latency in GPU(s)           | 4.92   | 2.68  | 0.66    | 1.42     | <b>1.09</b>  |
| Latency in CPU(s)           | 176.23 | 94.79 | 24.77   | 78.56    | <b>62.32</b> |

## V. DISCUSSION

Table I and Table II demonstrate that our model can accurately locate J-peaks in the test set with lower latency and fewer parameters. In detail, RPnet and U-Net can detect J-peaks accurately. However, due to the high latency and the large memory size, these two models are not suitable under practical implementation. When the number of convolutional layers of U-Net decreases to 8, the recall of the efficient U-Net-8 is less than 0.9, resulting in a poor distinction between J-peaks and adjacent peaks even with good data quality. It is worth noting that MIRD-Net can balance the accuracy and efficiency when detecting J-peaks. This is due to multi-level features are obtained by MIRD-Block. In addition, it can be concluded that detecting performance is further improved by the proposed model using the lightweight strategy. As shown in Fig.3, we observed that the network can still accurately locate J-peak even under slight artifact movement, which demonstrates the robustness of our model.

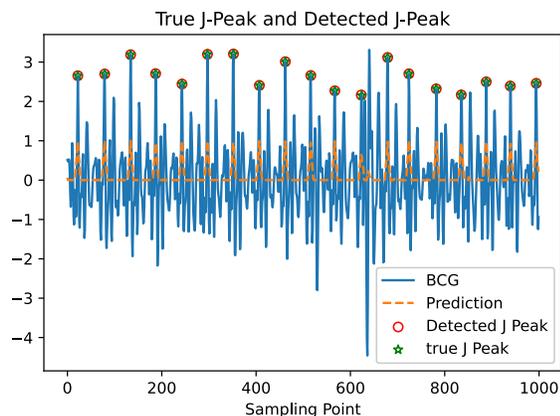


Fig. 3. An example of accurately locating J-peak in slight body movements.

## VI. CONCLUSION

To the best of our knowledge, no study has ever used a lightweight deep learning framework for the J-peak detection of BCG. We take the advantage of fewer parameters and higher segmentation accuracy of MIRD-Net to detect J-peak efficiently. Based on MIRD-Net, we replace the

standard convolution with lightweight depthwise separable convolution, add inverted residual blocks and remove the non-linearities in the narrow layers. In order to reduce the influence of individual differences on J-peak detection, we use an elegant transformation method to simulate the J-wave. This detection method achieves a high F1 score in new individuals and can detect J-peak under slight body movement. In the future, our research will focus on BCG J-peak detection of patients with cardiovascular disease, and our lightweight strategy will be applied to other neural network structures of the peak-detection task.

## REFERENCES

- [1] E. Pinheiro, O. Postolache, and P. Girão, "Theory and developments in an unobtrusive cardiovascular system representation: ballistocardiography," *The open biomedical engineering journal*, vol. 4, p. 201, 2010.
- [2] Kim, Chang-Sei and Ober, Stephanie L and McMurtry, M Sean and Finegan, Barry A and Inan, Omer T and Mukkamala, Ramakrishna and Hahn, Jin-Oh, "Ballistocardiogram: Mechanism and potential for unobtrusive cardiovascular health monitoring," *Scientific reports*, vol. 6, p. 31297, 2016.
- [3] Kim, Chang-Sei and Carek, Andrew M and Inan, Omer T and Mukkamala, Ramakrishna and Hahn, Jin-Oh, "Ballistocardiogram-based approach to cuffless blood pressure monitoring: proof of concept and potential challenges," in *IEEE Transactions on Biomedical Engineering*, 2018, pp. 2384-2391.
- [4] V. Moskalenko, N. Zolotykh, and G. Osipov, "Deep learning for ecg segmentation," in *International Conference on Neuroinformatics*. Springer, 2019, pp. 246-254.
- [5] G. Cathelain, B. Rivet, S. Achard, J. Bergounioux, and F. Jouen, "U-net neural network for heartbeat detection in ballistocardiography," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*, 2020, pp. 465-468.
- [6] S. Vijayarangan, V. R., et al., "Rpnnet: A deep learning approach for robust r peak detection in noisy ecg," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*, 2020, pp. 345-348.
- [7] H. Abrishami, C. Han, X. Zhou, M. Campbell, and R. Czosek, "Supervised ecg interval segmentation using lstm neural network," in *Proceedings of the International Conference on Bioinformatics & Computational Biology (BIOCOMP)*. The Steering Committee of The World Congress in Computer Science, 2018, pp. 71-77.
- [8] D. Hai, C. Chen, R. Yi, S. Gou, B. Yu Su, C. Jiao, and M. Skubic, "Heartbeat detection and rate estimation from ballistocardiograms using the gated recurrent unit network," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*, 2020, pp. 451-454.
- [9] Y. Huang, X. Li, C. Yan, L. Liu, and H. Dai, "Mird-net for medical-image segmentation," in *Advances in Knowledge Discovery and Data Mining*. Cham: Springer International Publishing, 2020, pp. 207-219.
- [10] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *arXiv preprint arXiv:1704.04861*, 2017.
- [11] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 4510-4520.
- [12] Y. Huang, C. Sun, T. Jin, S. Yang, and Z. Zhang, "Unobtrusive inter-beat interval estimation from multichannel ballistocardiogram signal using kalman filter," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*, 2020, pp. 455-460.
- [13] Alivar, Alaleh, et al. "Motion artifact detection and reduction in bed-based ballistocardiogram." *IEEE Access*, 2019, pp. 13693-13703.
- [14] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
- [15] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234-241.