# LETTER High Noise Tolerant R-Peak Detection Method Based on Deep Convolution Neural Network\*

Menghan JIA<sup>†a)</sup>, Feiteng LI<sup>†</sup>, Zhijian CHEN<sup>†</sup>, Nonmembers, Xiaoyan XIANG<sup>††</sup>, Member, and Xiaolang YAN<sup>†</sup>, Nonmember

**SUMMARY** An R-peak detection method with a high noise tolerance is presented in this paper. This method utilizes a customized deep convolution neural network (DCNN) to extract morphological and temporal features from sliced electrocardiogram (ECG) signals. The proposed network adopts multiple parallel dilated convolution layers to analyze features from diverse fields of view. A sliding window slices the original ECG signals into segments, and then the network calculates one segment at a time and outputs every point's probability of belonging to the R-peak regions. After a binarization and a deburring operation, the occurrence time of the Rpeaks can be located. Experimental results based on the MIT-BIH database show that the R-peak detection accuracies can be significantly improved under high intensity of the electrode motion artifact or muscle artifact noise, which reveals a higher performance than state-of-the-art methods.

*key words:* electrocardiograph, noise tolerance, convolution neural network, dilated convolution

#### 1. Introduction

Growing aging population and aggravating environmental problems have led to an increasing interest in wearable healthcare monitoring devices. Being an important part, the wearable electrocardiogram (ECG) monitoring system is widely concerned. In general, the ECG data is sampled at wearable devices and further analyzed on the cloud. However, with complex environments of the acquisition devices, the sampled ECG signals are often superimposed with various kinds of noises, such as power line interference (PLI), baseline wandering with respiration (BW), electrode motion artifact (EM) and muscle artifact (MA) noise. Among them, the BW and EM noises can reach 100% peak-to-peak ECG amplitude, while the MA noise can even reach 500% during a period of 100-500ms [1], which results in a great challenge to the analysis of ECG signals.

Researchers have proposed several noise tolerant algorithms to analyze ECG signals and detect R-peaks [1]– [5]. Friesen G M et al [1] utilizes the amplitude, derivative, and frequency spectrum of ECG signals to identify Rpeaks. [2]–[4] change the shape of the signal by mathe-

Manuscript revised July 16, 2019.

Manuscript publicized August 2, 2019.

<sup>†</sup>The authors are with the Institute of VLSI Design, Zhejiang University, Hangzhou, China.

a) E-mail: jiamengh@sina.com

DOI: 10.1587/transinf.2019EDL8097

matical transformations and then locate R-peaks based on the amplitude and other features of the ECG. To improve the robustness, both [3] and [4] filter the signal first. However, filtering does not work for the EM or MA noise, because the spectrum distributions of these noises are similar to that of the effective ECG signals. Nakai Y et al [5] adds a template matching technique with short time autocorrection on the traditional detection method such as continuous wavelet transform (CWT). Experiments in [5] show that with the increase of noise intensity, its detection performance drops greatly. Therefore, in extreme environments with high-intensity noises, all the methods above are not suitable for R-peak detection.

Different from finding features manually, in this letter, we propose a noise tolerant R-peak detecting algorithm based on a deep convolution neural network. The network inputs a piece of sliced ECG signal at a time and gives every point's probability of belonging to the R-peak regions. Then a post-processing step is added to analyze the probabilities and get the location of R-peaks. We evaluate this method by the MIT-BIH database [6]–[8] and experimental results show that the R-peak detection performance of this method can be significantly improved at low SNR.

# 2. Methods

There are certain morphological and temporal features of Rpeaks in the ECG. For example, R-peaks tend to have larger amplitude, and the intervals of consecutive R-peaks are similar. In the case of high-intensity noises, some morphological features can still be retained, and the temporal features of R-peak will not be destroyed. Therefore, we use a deep learning network to learn the morphological and temporal features of R-peaks in noisy ECG signals. We consider the sampled ECG points within 0.05s from R-peaks as the Rpeak regions. The points in the R-peak regions are classified as "1", and the other points are classified as "0". The network outputs the predicted probabilities. Then all the network outputs are analyzed in the post-processing stage to locate the R-peaks.

### 2.1 Network Architecture

The blocking diagram of the proposed deep convolution neural network (DCNN) is shown in Fig. 1. The first five layers of convolution are used for preliminary feature ex-

Manuscript received May 16, 2019.

<sup>&</sup>lt;sup>††</sup>The author is with the Department of Microelectronics, Fudan University, Shanghai, China.

<sup>\*</sup>This paper is partilally supported by National Natural Science Foundation of China under Grant 61801425.



Fig. 1 Blocking diagram of the proposed R-peak detection method. Where "Conv" stands for the convolution layer, "BN" stands for the batchnormalization layer, "ReLU" stands for the activation layer with rectified linear units and "Dilated Conv" stands for the dilated convolution layer.

traction from the data. Then, two parallel dilated convolution layers with different sampling rates, also called atrous spatial pyramid pooling (ASPP) layers in [9] are added. Because of the diversity of the convolution scales, the network can extract morphological and temporal features from diverse fields of view. The last convolution layer and the softmax layer are used to comprehensively analyze all the features and get the predicted probabilities. After each convolution computation, we add a batch-normalization (BN) layer to get higher learning rates and be less careful about the initialization of the weights [10]. Rectified linear units (ReLU) are used for activation function to train faster while keeping a high performance [11]. The original ECG recording contains the most morphological features, so we slice pieces of continuous data from it as the inputs of the network by a sliding window. If the recording contains multiple channels, all the channels will input to the network simultaneously, and the network will fuse different channels and output a comprehensive prediction. After that, the window moves a certain distance in the direction of time and slices another data as the next input. By shortening the moving distance of the sliding window, for example, setting the moving distance to half the width of the sliding window itself, every sample point can be analyzed multi times for higher detection performance, as shown in Fig. 2. The average value of the input data is set to 0 to make it easier for the network to extract features between data without being affected by the offset.



**Fig. 2** The sliding window extracts a piece of continuous data from the original recording and moves half of the width at a time. The example ECG data is from record 100 of the MIT-BIH Arrhythmia Database (MITDB) added electrode motion noise at 6dB SNR. The sampling frequency of the MITDB is 360Hz. Therefore, the width of the sliding window in the figure is 360 points and the window moves 180 points at a time.



**Fig.3** Operation of post-process from four consecutive network outputs. The input data of the network is shown in Fig. 2

#### 2.2 Post-Process

The schematic diagram of the post-process in this method is shown in Fig. 3. When completing the scanning of one ECG recording by the sliding window, the network analyzes multiple times for each sampling point and yields multiple probabilities. We average the probabilities of the same point for higher performance and then binarize the averaged probabilities. There are two kinds of burrs in the binarized signal. The first kind of burrs is the "0" values in continuous "1"s and the second kind of burrs is the "1" values in continuous "0"s, as shown in the dotted boxes of Fig. 3. Therefore, we add a deburring step to remove these burrs. First, we consider the continuous "0" values with widths less than 0.01s as the first kind of burrs and correct them to "1"s. Then the continuous "1" values with widths less than 0.03s are considered as the second kind of burrs and we correct them to "0"s. Finally, the mid-point of continuous "1" values is considered to be the occurrence time of the R-peak.

#### 3. Experiment Results

The classification performance of the proposed method is evaluated by the public ECG signal database, MIT-BIH Arrhythmia Database (MITDB) [6], whose sample frequency is 360 and the noise database, MIT-BIH Noise Stress Test database (NSTDB) [7]. We select for the performance evaluation of our technique with the BW, EM, and MA noises, as recommended in the NSTDB. Three types of the noises mentioned above with several SNRs are added for training and testing by using the tool "nst" provided by the PhysioNet [7], which is a necessary tool for [7]. Same as [12], 22 recordings carefully selected in [6] are used for training and another 22 for testing. As a widely accepted classification, it ensures that the number of R-peaks and beat types in each dataset are similar. During the training stage, we use the ECG signals contaminated by different noises with different SNRs (set as 24, 12, 6, 0) as the inputs. The proportions of the training data containing EM, MA or BW noise are all  $\frac{1}{3}$ , making the noise tolerant abilities of the network similar.

The performance of the R-peak detection is measured by the sensitivity (*S e*) and positive predictivity (<sup>+</sup>*P*), as recommended by the American Standard ANSI/AAMI EC57 [13]. They are defined as Eq. (1), where *TP*, *FN*, and *FP* are the number of true-positive, false-negative and false-positive events, respectively. Comparisons with the proposed method and the continuous wavelet transform (CWT) [2], Hilbert transform with second threshold (HT w/ 2nd Th) [3], empirical mode decomposition (EMD) [4] and short-term autocorrelation with template (STAC w/ TM) [5] for noise tolerant R-peak detection performances are shown in Fig. 4, 5, 6. In addition, there are only R-peak detection performances with the EM and MA noises in [2], [5].

$$S_e = \frac{TP}{TP + FN}$$

$$^+P = \frac{TP}{TP + FP}$$
(1)

The proposed method achieves 99.72% *S e* and 99.76% <sup>+</sup>*P* when there is no noise in the test sets, reaching a relatively high level. Due to the excellent feature extraction and analysis capabilities, this method gets a significantly higher R-peak detecting performance than traditional methods for the ECG signals contaminated by the EM or MA noise. The *S e* and <sup>+</sup>*P* of [2]–[5] all degrade substantially as the EM or MA noise increases because these two noises fluctuate frequently in a short time and traditional detection methods have been deceived. For example, when the SNR decreased to 0dB under the EM noise, the *S e* and <sup>+</sup>*P* of this method are 96.7% and 94.1%, 5.6% and 22.5% higher than [3] and much higher than others, as shown in Fig. 4. When



**Fig.4** The performance comparisons for the ECG contaminated by the EM noise at different SNRs. (a) The sensitivity comparison. (b) The positive predictivity comparison.



**Fig.5** The performances comparison for the ECG contaminated by the MA noise at different SNRs. (a) The sensitivity comparison. (b) The positive predictivity comparison.



**Fig.6** The performance comparisons for the ECG contaminated by the BW noise at different SNRs. (a) The sensitivity comparison. (b) The positive predictivity comparison.

the ECG data is contaminated by the MA noise, the performance of this method is also much higher than that of traditional methods, as shown in Fig. 5.

The frequency spectrum distribution of the BW noise is different from that of most effective ECG signals [1], therefore, traditional methods can remove the BW noise at the expense of only a small part of morphological features and achieve high detecting performance. To balance the robustness against all noises, the proposed method retains all the morphological features of the ECG signals. When training, the network will automatically get the best total noise tolerant performance. However, at this moment, the noise tolerant ability of the network against the BW noise is slightly lower than other methods, as shown in Fig. 6. Even so, the proposed algorithm can still achieve 92.3% *S e* and 92.0%  $^+P$  scores at -6dB, making it possible to work well under higher intensity of the BW noise.

#### 4. Conclusion

This letter proposes a noise tolerant algorithm to detect Rpeaks. The proposed algorithm is combined with a deep convolution neural network to give the probability of which sampling points belonging to the R-peak regions and a postprocessing step to locate R-peaks. Experiments based on the MIT-BIH database reveal that the R-peak detection accuracies can be significantly improved under high-intensity noises. Even at 0dB SNR, the proposed method can also get more than 90% *S e* and <sup>+</sup>*P* scores under the BW, EM or MA noise.

## Acknowledgments

The authors acknowledge the support of the National Natural Science Foundation of China under Grant 61801425.

#### References

- G.M. Friesen, T.C. Jannett, M.A. Jadallah, S.L. Yates, S.R. Quint, and H.T. Nagle, "A comparison of the noise sensitivity of nine QRS detection algorithms," IEEE Trans. Biomed. Eng., vol.37, no.1, pp.85–98, Jan. 1990.
- [2] I.R. Legarreta, P.S. Addison, M.J. Reed, N. Grubb, G.R. Clegg, C.E. Robertson, and J.N. Watson, "Continuous Wavelet Transform Modulus Maxima Analysis of the Electrocardiogram: Beat Characterisation and Beat-to-beat Measurement," Int. J. Wavelets Multiresolut Inf. Process., vol.03, no.01, pp.19–42, March 2005.
- [3] N.M. Arzeno, Z.-D. Deng, and C.-S. Poon, "Analysis of First-Derivative Based QRS Detection Algorithms," IEEE Trans. Biomed. Eng., vol.55, no.2, pp.478–484, Feb. 2008.

- [4] H. Li, X. Wang, L. Chen, and E. Li, "Denoising and R-Peak Detection of Electrocardiogram Signal Based on EMD and Improved Approximate Envelope," Circuits, Systems, and Signal Processing, vol.33, no.4, pp.1261–1276, April 2014.
- [5] Y. Nakai, S. Izumi, M. Nakano, K. Yamashita, T. Fujii, H. Kawaguchi, and M. Yoshimoto, "Noise tolerant QRS detection using template matching with short-term autocorrelation," 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Chicago, IL, pp.34–37, IEEE, Aug. 2014.
- [6] G.B. Moody and R.G. Mark, "The impact of the MIT-BIH Arrhythmia Database," IEEE Eng. Med. Biol. Mag., vol.20, no.3, pp.45–50, June 2001.
- [7] G.B. Moody, W. Muldrow, and R.G. Mark, "A noise stress test for arrhythmia detectors," Computers in Cardiology, vol.11, no.3, pp.381– 384, 1984.
- [8] A.L. Goldberger, L.A.N. Amaral, L. Glass, J.M. Hausdorff, P.C. Ivanov, R.G. Mark, J.E. Mietus, G.B. Moody, C.-K. Peng, and H.E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals," Circulation, vol.101, no.23, June 2000.
- [9] L.C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A.L. Yuille, "DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs," arXiv:1606.00915 [cs], June 2016.
- [10] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," Proceedings of the 32nd International Conference on Machine Learning - Volume 37, ICML'15, pp.448–456, 2015.
- [11] V. Nair and G.E. Hinton, "Rectified linear units improve restricted boltzmann machines," Proceedings of the 27th International Conference on Machine Learning (ICML-10), pp.807–814, 2010.
- [12] P. deChazal, M. O'Dwyer, and R.B. Reilly, "Automatic Classification of Heartbeats Using ECG Morphology and Heartbeat Interval Features," IEEE Trans. Biomed. Eng., vol.51, no.7, pp.1196–1206, July 2004.
- [13] Testing and Reporting Performance Results of Cardiac Rhythm and ST Segment Measurement Algorithms, Association for the Advancement of Medical Instrumentation, 1998.