R-PEAK DETECTION FROM NOISY ECG DATA USING MULTI-CHANNEL 1D-CNN WITH ACCELEROMETER INPUT

Tetsuya Hirota
Ryugo Fujita
Atsushi Harada
Daisuke Kawamura
Tokai Rika Co., Ltd.
Japan

Keiichi Yamada
Meijo Univ.,
Japan

Paper Number 23-0047

ABSTRACT

In order to prevent traffic accidents due to abrupt changes in the driver's health condition, we have proposed a non-contact type electrocardiographic sensor that monitors the electrocardiogram (ECG) of a driver holding a steering wheel while seated. However, the heart rate detection accuracy degrades while driving due to the lower signal-to-noise ratio (SNR) of the ECG caused by the noise from vehicle vibration and static electricity, among others. In this study, we propose a method of detecting R-peaks of the ECG from the low SNR ECG signal with high accuracy using a multi-channel one-dimensional convolutional neural network with accelerometer signals as an input. As the results, we achieved an *F*-score of 78.5% and a root-mean-square error (RMSE) of 1.99 ms. The R-peak detection performance was significantly improved when the input data length of around 1100 ms was chosen.

1. INTRODUCTION

Today the number of car accidents still remains at a high level. It is believed that many cases of the accidents are attributable to human errors such as carelessness of the driver or violation of the Road Traffic Law; on the other hand, there are not a few cases attributable to abrupt changes in the health condition of the driver, which was caused by his/her underlying illness.

In recent times, heart diseases are the leading cause of death [1], which accounts for a large percentage of car accidents caused by driver illness [2]. Because the risk of heart diseases increases exponentially with age, considering that many countries will face an aging society and the number of elderly drivers is expected to increase in the future, the number of traffic accidents caused by heart diseases is expected to increase.

Therefore, it is one of the urgent tasks to develop a system that can detect cardiovascular abnormalities that arise during driving a vehicle by monitoring the heart activities such as the ECG or heart rate of the driver, and carry out appropriate driving interventions, such as moving the vehicle to the shoulder and safely stopping it, and take appropriate measures to rescue the driver, such as notifying the abnormality of the driver to other vehicles or calling of the emergency services.

The purpose of this study is to accurately detect the driver's heart rate interval (R to R interval, RRI) by using a non-contact method without the need to attach electrodes to the body surface. For heart rate variability (HRV) analysis [3], RRI should be acquired with high accuracy of about several milliseconds. A capacitively coupled electrocardiographic (cECG) sensor is one of the typical methods for monitoring the heartbeat in a contactless manner. However, it is difficult to obtain RRI with high enough accuracy with the method while driving, because noise caused by vehicle vibrations etc. will superimpose the cECG signals to decrease the signal-to-noise ratio (SNR) [4].

In this study, we propose a method to accurately detect the driver's RRIs in a moving vehicle by using the accelerometer signal of the vehicle together with the cECG signal. A one-dimensional convolutional neural network (1D-CNN) using cECG signals and accelerometer signals as multichannel inputs is used to detect the R-peaks for acquiring RRI.

The proposed method is evaluated by experiments using data acquired from 4 subjects while driving and the effectiveness of the method is demonstrated. The method detects R-peaks in the low SNR cECG signal with a *F*-score of 78.5% when the input window length is 1100 ms.

2. RELATED WORKS

Methods for monitoring the activity of the heart fall roughly into two categories: contact type methods and non-contact type methods.

In the contact type methods, the potential differences between two or more body surfaces sandwiching the heart are obtained by attaching the electrodes to the surfaces, but this forces the driver to attach the electrodes every time he or she gets on the vehicle, and therefore is unrealistic to apply to driver monitoring.

On the other hand, the non-contact type method does not have such disadvantages, but has a problem that the signal quality is unstable because the signal is easily affected by temperature, humidity, body movement or static electricity.

Non-contact monitoring of heart activity includes electrocardiogram monitoring using a cECG sensor [5], ballistocardiogram monitoring [6], magnetocardiography monitoring using a magnetic impedance sensor [7], and Doppler sensing [8], and several experiments of monitoring heart activity in running vehicles using these methods have been reported [4] [9] [10]. Among them, a cECG sensor is relatively resistant to noise as compared with other non-contact type sensors.

ECG waveform of one heartbeat is composed of five consecutive waves, P, Q, R, S, and T waves, as shown in Figure 1. An adaptive correlation filter [11] can be used to detect signals such as QRS signals from noisy ECG data. However, the detection fails if the noise intensity becomes greater than the signal intensity.

R-peak detection from noise-intensive ECG acquired in a running vehicle using CNN are reported [12] [13] [14], but the precision level required for HRV analysis has not been attained.

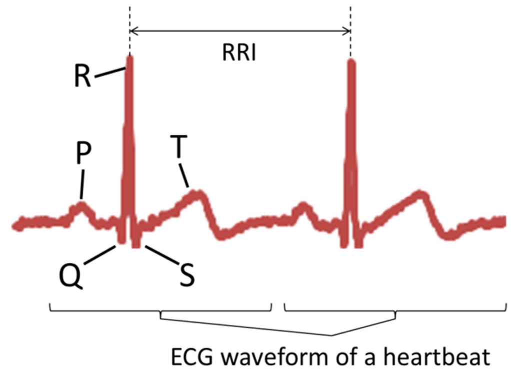


Figure 1. A typical ECG waveform. An ECG waveform of one heartbeat is composed of five consecutive waves, P, Q, R, S, and T waves.

3. PROPOSED METHOD

The architecture of the proposed method is shown in Figure 2.

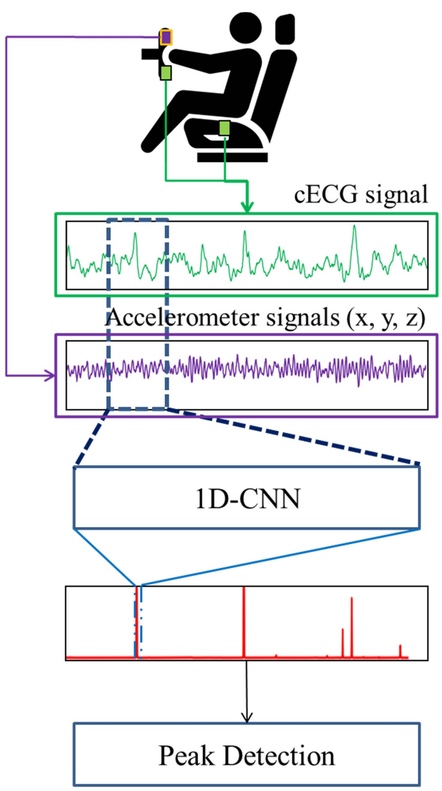


Figure 2. Schematic overview of the proposed method

We propose a 1D-CNN model for inferring the probability of the presence of the R-peak of an ECG at each moment from low SNR cECG signal and accelerometer signals acquired synchronously.

Considering that the low SNR of the cECG signal is due to the noise caused by vehicle vibrations, accelerometer signals, which are considered to be correlated with vehicle vibrations, are used as inputs of the model to remove the noise.

Given that an ECG waveform is composed of five consecutive waves, P, Q, R, S, and T waves (see Figure 1), in order to detect R-peaks, inclusion of the information of P, Q, S, and T waves within the same beat as the R-peak would be effective. Therefore, 1D-CNN layers are adopted because they are able to incorporate local time series relations of inputs.

The input to the 1D-CNN is the amplitude data of the cECG and accelerometer signals for a specific duration, and the output is the existence probability of the R-peak at the center of the duration. The R-peak timing is inferred by detecting the moment when the probability exceeds a threshold and detecting the local maximum of the probability.

4. EXPERIMENT

4-1 Data

We evaluated the proposed method with cECG signals acquired using a cECG system integrated into the passenger seat of a car [4] and acceleration signals acquired using a three-axis accelerometer attached to the steering wheel. For the reference signal, a contact-type ECG sensor (NeXus-10 Mark II, a multi-sensor physiological measurement system made by Mind Media Co.) with adhesive electrodes was used.

We acquired cECG signals and reference ECG signals of 6 subjects seated on the passenger seat of a running vehicle, and used data from 4 subjects whose reference ECG signals were measured with sufficient intensity. The total length of the data from the 4 subjects was 20 minutes.

The sampling rate was 1000 Hz, 2048 Hz and 1000 Hz for the cECG signal, reference contact-type ECG signal, and the accelerometer signal, respectively, and the contact-type ECG signal was subsequently downsampled to 1000 Hz.

4-2 Model Overview

The structure of the 1D-CNN model used in the experiment is shown in Figure 3. The input of the model was the cECG and accelerometer signals for the duration of 500 ms (500 points in total, since sampling is performed at 1000 Hz), and the output was the existence probability of the R-peak at the moment 250 ms from the beginning.

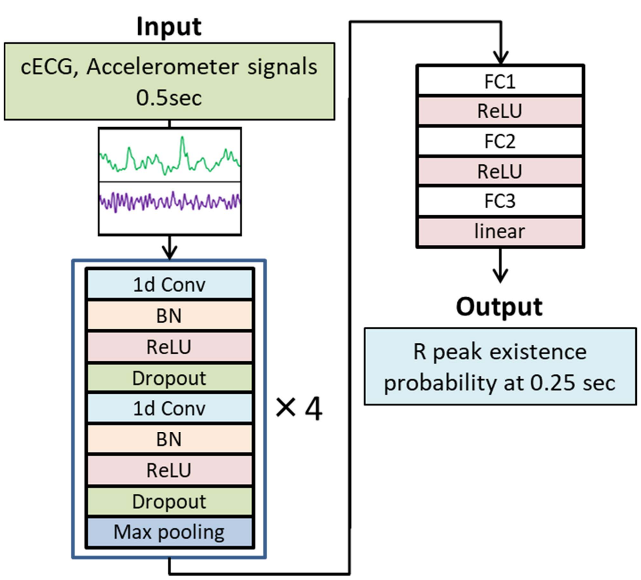


Table 1. Parameters of the network

Layer	Parameter
1dConv	kernel size:16 filters:64
Dropout	0.2
Max pooling	2
FC1	output units:16
FC2	output units:4
FC3	output units:1

Figure 3. The architecture of the network

Each convolution block consisted of two 1-dimensional convolutional layers (1d Conv), two dropout layers, two batch normalization layers (BN), two activation functions (ReLU), and one pooling layer (Max Pooling). After repeating the block four times, three fully connected layers (FC) were applied. The parameters for each layer of the network are shown in Table 1. All the layers with the same function (layer name) had the same parameter values.

4-3 Training

In training the 1D-CNN model, 500 ms-length cECG and accelerometer signal sections were extracted and used as inputs, and R-peak labels created from the reference signals were used as annotations. Specifically, if an R-peak existed at the center of the 500 ms reference signal section, the annotation was 1, otherwise 0.

The training data set was created by extracting an input/annotation pair from cECG, accelerometer and reference signals, and sliding the whole data by 1 ms (the sampling interval) to extract another pair, and so forth, and then the model training was performed using the data set.

4-4 Evaluation Method

The output from the trained 1D-CNN model, which takes on a fractional value, was thresholded to give either 0 (R-peak not found) or 1 (R-peak found) as the final output, and the performance was evaluated. For evaluation met-

$$Precision = \frac{TP}{TP + FP} \times 100$$

$$Recall = \frac{TP}{TP + FN} \times 100$$

$$F\ score = \frac{2 \times Precision \times Recall}{Precision + Recall} \times 100$$

rics, Precision, Recall, and *F*-score were used. Each metric was calculated as follows:

The definitions of TP, FP and FN are as follows:

TP : The number of R-peaks which was detected within the tolerance window of 20 ms of a true R-peak, and for which no other R-peaks are detected within the window. The true R-peaks are obtained from the reference signal.

FP : The number of R-peaks which was detected outside the tolerance window mentioned above, or for which other instances of such R-peaks are also detected within the same window (each detected instance will be counted).

FN : The number of true R-peaks which was not detected.

In addition, all R-peaks detected within the tolerance windows were evaluated in terms of the temporal root-mean-square error (RMSE) from the true R-peak locations.

Leave-one-subject-out cross validation was performed on the data and the average of the evaluation metrics over the 4 cross-validation evaluations was used as the R-peak detection performance of the model. In each of 4 cross-validation evaluations, data for two subjects, one subject, and one subject were used for the training data, the validation data, and the test data, respectively.

5. RESULTS AND DISCUSSION

5-1 Choice of the Accelerometer Channels to Use

The choice of the input channels of the accelerometer to be used can have an impact on the performance of the model. Here such choice was studied. The input channel candidates were cECG and accelerometer signals in three directions (Acc_x, Acc_y, Acc_z) as shown in Figure 4 obtained from the accelerometer.

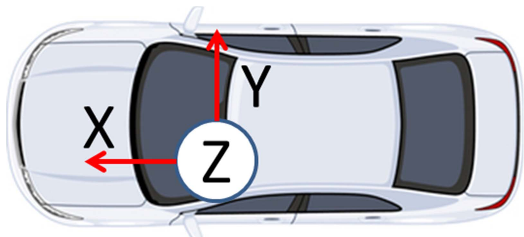


Figure 4. Three axis directions of the accelerometer

Table 2. Experimental results. Performance of R-peak detection on cECG single depending on the input signals to the 1D-CNN.

Input	Recall	Precision	*F*-score	RMSE
cECG	70.5	69.0	69.7	2.01
cECG, Acc_x	74.8	69.6	72.1	2.16
cECG, Acc_y	77.4	67.5	72.1	2.01
cECG, Acc_z	72.8	70.2	71.4	2.19
cECG, Acc_x, Acc_y, Acc_z	70.1	76.2	73.0	2.14

The results are shown in Table 2.

When all of Acc_x, Acc_y and Acc_z are used as input, the *F*-score is larger by 3% or more compared to the case where only cECG is used, and the RMSE is also a sufficiently small value of about 2 ms. Therefore, it can be said that the use of the accelerometer signals is advantageous for detecting R-peaks.

Figure 5 shows the waveforms of cECG, reference ECG, the vertical component of the accelerometer signal (Acc_z) and the output of the model.

In this example, a correlation is seen between the cECG and the accelerometer signal: where the noise intensity is large in cECG, the amplitude of the accelerometer signal is also large.

It is considered that the influence of the noise superimposed on the ECG signal can be canceled by using the accelerometer signal as an input of the 1D-CNN model, and as a result, erroneous R-peak detections can be suppressed.

5-2 Choice of the Input Data Length

The length of the data section will affect the performance of the model, and here such effect was studied. 1D-CNN model training was performed by choosing a different value for the input data section length, from 100 ms to 1500 ms. Note that the model inputs were cECG, Acc_x, Acc_y and Acc_z. The model output was the existence probability of the R-peak at the center of the input data section (for example, the output is the existence probability of the R-peak at 750 ms from the beginning of the data section if the input data length is 1500 ms), which was then thresholded to give either 0 or 1 as mentioned above. The results are shown in Table 3.

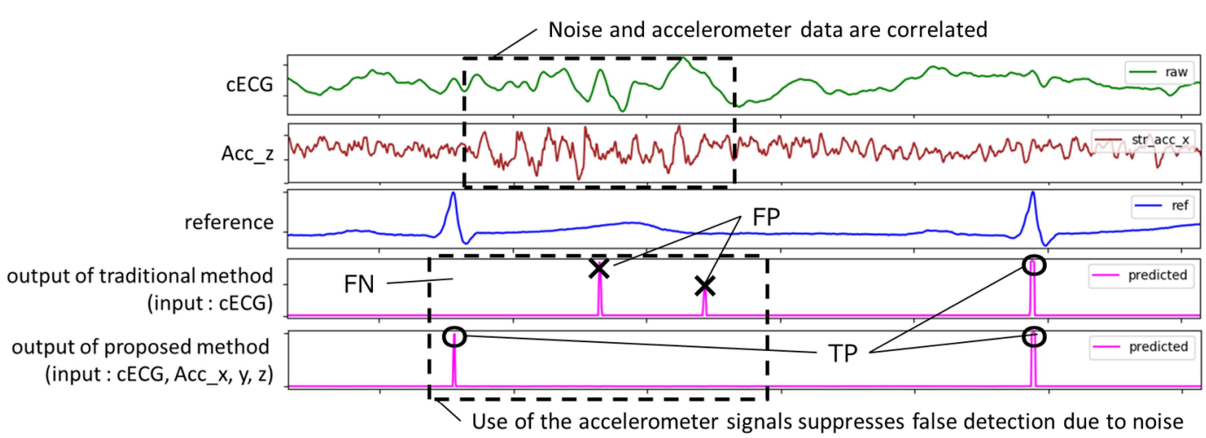
Generally, as the input data section length increases, the detection performance improves. The maximum *F*-score is achieved when the input data section length is 1100 ms, and thereafter, it starts to decrease. This may be due to the following reasons:

The RRIs of a resting healthy human are about 600 ms to 1000 ms, and the RRIs of our subjects are also roughly distributed in that range. Therefore, in the data used here, there are no other R-peaks in the range of around 1200 ms centered at one of the R-peaks. When the input data section length is 1100 ms (< 600 ms × 2), if an R-peak is located at the center of the data section (= output is 1), no other R-peaks exist in the interval. However, if the R-peak deviates from the center by several tens of ms or more, other R-peaks will enter the section. In other words, when the input data length is around 1100 ms, in addition to whether the R-peak is actually found at the center of the section (Criterion 1), whether no other R-peaks are found in the section (Criterion 2) can be used to determine whether we do have an R-peak at the center of the interval, so hence improving the accuracy.

If the input data section length is too short, the detection performance deteriorates due to insufficient information on the neighbor waves (P, Q, S, and T waves) in the input section, in which case the Criterion 1 becomes unreliable; and if the input data section length is too long (> 600 ms × 2), even if an R-peak is located at the center, other R-peaks can enter the input data section, and therefore the Criterion 2 becomes irrelevant (it won't be used). In either case, the detection performance is expected to deteriorate.

Table 3. Performance of R-peak detection depending on the input data window length (100 ~ 1500 ms).

Input data length	Recall	Precision	*F*-score	RMSE
100 ms	66.5	66.3	66.4	2.02
300 ms	71.6	71.8	71.7	2.26
500 ms	70.1	76.2	73.0	2.14
700 ms	72.7	73.4	73.0	1.98
900 ms	76.6	74.4	75.5	2.24
1100 ms	75.7	81.5	78.5	1.99
1300 ms	74.4	72.6	73.5	2.00
1500 ms	71.7	73.1	72.4	2.22



6. CONCLUSION

In this study, we proposed a method to accurately detect the driver's RRI in a running vehicle by the 1D-CNN, using the multi-channel inputs consisting of the accelerometer signals of the vehicle and the cECG of the driver, and discussed its results.

Figure 5. Improvement of detection performance by adding accelerometer inputs

We confirmed that the detection performance was improved by more than 3% points in *F*-score by using the accelerometer signals as an input to the 1D-CNN together with the cECG data, acquired for a total of 20 minutes for 4 subjects while driving. In addition, the detection performance was improved as the input data section length was increased, and the maximum *F*-score of 78.5% is achieved when the input data section length is 1100 ms. Furthermore, under all conditions, a sufficiently small RMSE of about 2 ms was achieved, and the R-peaks were detected with sufficient accuracy to withstand HRV analysis.

As a future work, we plan to add more training data to handle ECGs with diverse characteristics. We also plan to implement features to detect specific health problems, such as arrhythmias.

REFERENCES

[1] World Health Organization (2011). Global Atlas on cardiovascular disease prevention and control, viewed 3 Dec. 2019, < http://whqlibdoc.who.int/publications/2011/9789241564373_eng.pdf?ua=1>

[2] European Commission (2013). New Standards for Driving and Cardiovascular Diseases, viewed 3 Dec. 2019, <https://ec.europa.eu/transport/road_safety/sites/roadsafety/files/pdf/behavior/driving_and_cardiovascular_disease_final.pdf>

[3] U. Rajendra Acharya, K. Paul Joseph, N. Kannathal, C. Min Lim, Jasjit S. Suri (2006). Heart rate variability: a review, Med Bio Eng Comput, vol.44, pp.1031–1051.

[4] M. Kunugita, A. Harada, Y. Yamanaka, Y. Mizuno (2019). Study on noise generation mechanism from static electricity in capacitively coupled ECG sensor, 58th Annual Conference of Japanese Society for Medical and Biological Engineering, Okinawa, Japan.

[5] Lopez A, Richardson PC (1969). Capacitive Electrocardiographic and Bioelectric Electrodes, IEEE Transactions on Biomedical Engineering, BME-vol.16 (1), pp.99.

[6] Koivistoinen T, Junnila S, Va¨rri A, Ko¨o¨bi T (2004). A new method for measuring the ballistocardiogram using EMFi sensors in a normal chair. 26th IEEE EMBS conference, San Francisco, CA, USA.

[7] Guardo R, Charron G, Goussard Y, Savard P (1995). Contactless recordings of cardiac related thoracic conductivity changes. 17th IEEE EMBS conference, Montreal, Canada.

[8] Ichapurapu R, Jain S, John G, Monday T, Lie DYC, Banister R, Griswold J (2009). A 2.4 GHz non-contact biosensor system for continuous vital-signs monitoring.10th Annual IEEE wireless and microwave technology conference, Clearwater, FL, USA.

[9] M. Walter, B. Eilebrecht, T. Wartzek, S. Leonhardt (2011). The smart car seat: personalized monitoring of vital signs in automotive applications, Personal and Ubiquitous Computing archive, Vol.15 (7), pp.707-715.

[10] S. Leonhardt, A. Aleksandrowicz (2009). A Non-contact ECG monitoring for automotive application, 5th International workshop on wearable and implantable body sensor networks, Hong Kong, China.

[11] M. Kisohara, Y. Masuda, E. Yuda, J. Hayano (2019). Usefulness of Adaptive Correlation Filter for Detecting QRS Waves from Noisy Electrocardiograms. 2019 IEEE 1st Global Conference on Life Sciences and Technologies, Osaka, Japan.

[12] C. H. Antink, E. Breuer, D. U. Uguz, S. Leonhardt (2018). Signal-Level Fusion with Convolutional Neural Networks for Capacitively Coupled ECG in the Car, 2018 Computing in Cardiology Conference, Maastricht, Netherlands.

[13] T. Hirota, R. Fujita, M. Otake, Y. Nawa, K. Yamada, D. Kawamura (2019). R-wave detection from low S/N electrocardiogram using 1D convolutional neural network, 8^{th}-aim-mtg of special interested group on artificial intelligence in medicine, Yokohama, Japan.

[14] Moody GB, Moody B, Silva I (2014). Robust Detection of Heart Beats in Multimodal Data: The PhysioNet/Computing in Cardiology Challenge 2014, Computing in Cardiology 2014, vol.41, pp.549–552.