Preprocessing Method for Performance Enhancement in CNN-Based STEMI Detection From 12-Lead ECG

YEONGHYEON PARK1, IL DONG YUN1, (Member, IEEE), AND SI-HYUCK KANG2
1Department of Computer and Electronic Systems Engineering, Hankuk University of Foreign Studies, Yongin 17035, South Korea
2Cardiovascular Center, Internal Medicine, Seoul National University Bundang Hospital, Seongnam 13620, South Korea
Corresponding author: Il Dong Yun (yun@hufs.ac.kr)

This work was supported in part by the National Research Foundation through the Basic Science Research Program of the Ministry of Education, Science, and Technology, South Korea, under Grant 2017R1A2B4004503 (NRF), in part by the Declaration of Helsinki, in part by the Seoul National University Bundang Hospital through the Institutional Review Board, in 2019, under Grant B-1811-502-003, and in part by the Hankuk University of Foreign Studies Research Fund.

ABSTRACT ST elevation myocardial infarction (STEMI) is an acute life-threatening disease. It shows a high mortality risk when a patient is not timely treated within the golden time, prompt diagnosis with limited information such as electrocardiogram (ECG) is crucial. However, previous studies among physicians and paramedics have shown that the accuracy of STEMI diagnosis by the ECG is not sufficient. Thus, we propose a detecting algorithm based on a convolutional neural network (CNN) for detecting the STEMI on 12-lead ECG in order to support physicians, especially in an emergency room. We mostly focus on enhancing the detecting performance using a preprocessing technique. First, we reduce the noise of ECG using a notch filter and high-pass filter. We also segment pulses from ECG to focus on the ST segment. We use 96 normal and 179 STEMI records provided by Seoul National University Bundang Hospital (SNUBH) for the experiment. The sensitivity, specificity, and area under the curve (AUC) of the receiver operating characteristic (ROC) curve are increased from 0.685, 0.350, and 0.526 to 0.932, 0.896, and 0.943, respectively, depending on the preprocessing technique. As our result shows, the proposed method is effective to enhance STEMI detecting performance. Also, the proposed algorithm would be expected to help timely and the accurate diagnosis of STEMI in clinical practices.

INDEX TERMS Convolutional neural network, electrocardiogram, noise reduction, pulse segmentation, ST elevation myocardial infarction.

I. INTRODUCTION
Acute Myocardial Infarction (MI) is a major cause of death in developed countries. MI can be divided into ST Elevation Myocardial Infarction (STEMI) or Non-ST Elevation Myocardial Infarction (NSTEMI). The complete occlusion of the coronary artery is called STEMI and it is characterized by a sudden shut-down of blood flow caused by thrombus or embolism [1]. The territory of involved heart muscle is usually large, and thus STEMI is potentially life-threatening when treated inappropriately.

The Electrocardiogram (ECG) is consisted by 5 waves (P, Q, R, S, and T) as shown in Fig. 1. The Q, R, and S waves are named as QRS complex after grouping, and the area between S and T is called as ST-segment. In the ECG of STEMI, the ST-segment is more elevated than the normal. The above characteristic is used to diagnose STEMI by physicians. However, they look similar in the overall view as shown in Fig. 2. In a case of emergency, the ECG may include noise caused by the electrical noise from a power cable or the movement of the patient. These noises are called mains hum and baseline wandering. When the baseline wandering is included in the ECG, it can lead to confusion when distinguishing between normal and STEMI. For example, the normal ECG which has an increasing trend by baseline wandering can be considered as STEMI and vice versa. The sample case is shown in the red box of Fig. 3.
Current guidelines recommend diagnosing STEMI within 10 minutes after a patient with chest pain first shows up to medical practitioners including physicians and paramedics [2]. However, previous study shows that the interpretation of ECG for diagnosing STEMI is unsatisfactory even among physicians [3]. In the study above, 124 physicians interpret 4392 ECG records. Their sensitivity and specificity for detecting STEMI is 65% and 79% respectively. In case above, 35% of the patients will not be properly treated in time. Thus, we develop the STEMI detection model to support physicians in an emergency room for making fast and accurate decision.

Basically, our purpose is detecting STEMI via ECG especially in emergency situation. We consider the ECG probably has more noise in the emergency situation. We adopt the Convolutional Neural Network (CNN) to construct a STEMI detection model. We mostly focus on enhancing the performance of STEMI detection in the same CNN architecture via the preprocessing technique. We describe related works in Section II. Then we present our approach in Section III. We experimentally confirm the effectiveness of performance enhancing method in Section IV and conclude in Section V.

II. RELATED WORK

Before presenting the method to improve the STEMI detection performance, we describe related works and discuss with them. We deal with the related studies for detecting MI with lead selection viewpoint in Section II-A. In Section II-B, we also present the difficulty of using 12 lead ECG due to the own characteristic of each lead. We describe the reason for using CNN-based algorithm in Section II-C.

A. REASON FOR USING 12-LEAD

Previously, some studies effort to detect MI via diverse methods [4]–[18]. For summarizing those, it can be categorized depends on whether using all leads of ECG or not.

Kumar et al. [4], Acharya et al. [5], and Lui and Chow [17] use wavelet transform, CNN, and CNN combined with RNN respectively to detect MI. They achieved high performance of MI detection but their studies have the same limitation. Location of an MI is differentiated based on the leads that are changed heart disease [19]. The localizations of MI is shown in Table 1 [20]. However, only single lead is used in above studies. In those case, they have the probability to miss the other localizations of MI of unused lead.

Xiao et al. [6] use three leads of ECG to consider a diverse of MI localization but it is still short. Their study does not consider the all localization of MI because they select and use some leads.

There are some studies using 12-lead ECG for detecting MI, although their aim is not considering all localizations of MI or STEMI. Martin et al. [7] effort to detect MI by statistically analysis with 12-leads. The sensitivity of their study is 50.0% and the specificity is 97.0%. Their method miss just a few normal people, but they miss half of the MI patients. Thus their algorithm is not sufficient for using in the hospital because of low sensitivity.

Haraldsson et al. [8] decompose 12-lead ECG using Hermite basis functions and use the result coefficients as inputs with the Neural Network (NN). They achieve sensitivity of 63.3% and specificity of 85.0%. In a study by Chang et al. [9], the MI is detected using Hidden Markov Models (HMM) and Gaussian Mixture Models (GMM) with sensitivity of 79.0% and precision of 68.7% but their method is also insufficient to using in clinical practice. Above studies show the increasing trend of sensitivity but it is still insufficient for using in clinical practice.

### TABLE 1. The localizations of Myocardial Infarction [20]. Those localizations commonly appear with ST-segment elevation.

<table>
<thead>
<tr>
<th>Localization</th>
<th>Lead</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anterior MI</td>
<td>V1-V6</td>
</tr>
<tr>
<td>Septal MI</td>
<td>V1-V4, disappearance of septum Q in leads V5 and V6</td>
</tr>
<tr>
<td>Lateral MI</td>
<td>I, aVL, V5, V6</td>
</tr>
<tr>
<td>Inferior MI</td>
<td>II, aVF</td>
</tr>
<tr>
<td>Posterior MI</td>
<td>V7, V8, V9</td>
</tr>
<tr>
<td>Right Ventricle MI</td>
<td>V1, V4R</td>
</tr>
<tr>
<td>Atrial MI</td>
<td>PII in I, V5, V6</td>
</tr>
</tbody>
</table>

VOLUME 7, 2019
Additionally, one of the studies confirms using 12-lead is more effective for diagnosing MI than selecting and using some of the leads in recent study [10]. They experiment with 12-lead, limb lead, lead-I, lead-II, and lead-III respectively. In their experiment, the performance on 12-lead is higher than other lead selections. Because when the characteristic of MI appears in the unchosen lead, the MI diagnosis may not be successful. Thus, we use all the leads of ECG for detecting and diagnosing STEMI.

Baloglu et al. [18] already experiment to detect all localizations of MI with plausible all subclasses of MI. The above study is evidence that to detect all localizations, all 12-leads should be used. Based on the above reasons, we use all 12-leads for STEMI detection.

B. DIFFICULTY OF FINDING THE QRS COMPLEX

In Fig. 2, we already show the similarity of the normal and STEMI in the overall view. On the other hand, when looking closely as the pulse unit, the distinguishing is more better by the ST-segment. Thus, pulse unit view may be considered more effective to diagnosing STEMI with computer aided algorithm. Above consideration should be confirmed as comparing the effectiveness for distinguishing normal or STEMI between observing the whole ECG and each pulse of ECG. We present the characteristic of pulse for each lead. Fig. 4 and Fig. 5 show the QRS complex of normal and STEMI respectively.

Finding the time location of the QRS complex, which is the center of the pulse, helps to segment the pulse information. However, finding the QRS complex with the manual process is inefficient so we need an automated algorithm. We need to consider the different characteristic of each lead to construct the algorithm. For example, the lead-I, II, and aVL show the shape of the generally known ECG, but lead-III and aVR have upside-down form.

When using only one lead as a criterion for segmenting pulse, a single QRS complex detector may perform well. However, finding the QRS complex in multi-channel ECG with a single QRS complex detector is as difficult as unlocking several locks with one key. This problem can be solved by setting up individual QRS detectors for each channel. For example, the threshold-based algorithm, which Lahiri et al. [21] used, may not work properly because it considers only one lead. Taking above into consideration, covering the various form of ECG is difficult. For instance, if the ST-segment elevated over the QRS complex as shown in lead-II of Fig. 5, the QRS complex detector cannot detect correctly.

Thomas et al. [22] make an effort to segment the ECG automatically. They train the HMM to segment the ECG into base1, P, base2, QRS, T, and base3. The HMM can be used for segmentation continuously and automatically once it is trained. However, they prepare the dataset for training HMM through manual segmentation. The limitation of their study is a high cost of preparing the dataset. In conclusion, we need to develop a simple but reliable QRS complex detector while eliminating the obstacles of QRS complex detection.

C. REASON OF CNN-BASED ALGORITHM

Earlier studies show high performance in MI classification problem with various classification model such as k Nearest Neighbors (kNN) algorithm [11], [12], Support Vector
Machine (SVM) [13]–[15], and Decision Tree (DT) [16]. Since those algorithms run based on the extracted features from raw data, the performance may be different depending on the feature selection. Thus, the effort to find sophisticated feature is needed before using feature-based classification algorithm.

Other limitations of above algorithms are following. First, kNN algorithm does not need training time but requires lots of time in test procedure [23]. SVM and DT can decide highly faster as real-time although needs training time [24], [25]. Moreover, non-linearity can be solved when using SVM and DT but requires costs such as following. The SVM needs large space for mapping low-dimensional non-linear feature to high-dimensional linear feature and DT needs to find the optimal combination and pruning for ensemble model construction.

However, CNN reduces above cost. Unlike feature-based classifier, convolution layers learn how to extract feature to represent raw data well via back propagation in the training phase [26], [27]. The difficulty of solving non-linearity is also eased by using non-linear activation function such as Rectified Linear Unit (ReLU) in information propagation [28]. In addition, CNN can classify in real-time after training the parameters. Thus, we adopt the CNN to classify the ECG as normal or STEMI.

III. PROPOSED APPROACH

Enhancing the performance in the same STEMI detecting model gives efficiency. We adopt noise reduction procedure to improve the distinguishing ability. Also, We segment pulses from ECG with the same reason as above. We describe the detail of whole procedure as below.

A. NOISE REDUCTION

Few previous studies cover the noise reduction procedure. Comparing to the previous study, we observe better performance after noise reduction [5]. Because, as we describe in Section I, diagnosis from ECG is confused by the noise. Thus, we adopt noise reduction to enhance the performance of the STEMI detection. We deal with the 12-lead ECG for 11 seconds provided by Seoul National University Bundang Hospital (SNUBH). Each ECG has mains hum at 60Hz, as well as baseline wandering.

Our purpose is not only to reduce noise to refine information but also to detect STEMI. Applying only the noise reduction method is not expensive but accumulated in each procedure including noise reduction, QRS complex detection, and STEMI detection results in an inefficient heavier algorithm. The method that used in previous studies to reduce mains hum [29]–[31] and baseline wandering [32], [33] is a bit expensive for us. Thus, we adopt and use the simple noise reduction method as a notch filter and high-pass filter.

We firstly apply a notch filter to ECG in order to eliminate the mains hum. We remove not only the 60Hz components but also the 2nd to 4th partial harmonic components associated with 60Hz. In the global perspective of ECG, baseline wandering is similar to low-frequency signals. Thus, we adopt the high-pass filter with cutoff frequency of 1Hz as a simple function to remove baseline wandering. The high-pass filter leads to slight distortions at the front of ECG but we find the QRS complex correctly in the overall view. The whole procedure of the above descriptions is summarized in Fig. 6.

B. PULSE SEGMENTATION

The NN performs better when the classes of dataset are clearly distinguishable. However, Normal and STEMI has a similar characteristic in overall view so the distinction between the above two class is confusing. We intend for NN to distinguish normal and STEMI by ST-segments but unlike our intention, NN can be trained in the direction to distinguish heart rate probably. For guiding the training direction, we adopt pulse segmentation. For training NN, segmenting only important information such as pulses can prevent the unintended learning of unnecessary information.

We use the QRS complex as a criterion of pulse segmentation. We describe earlier that finding QRS complex may not be always fine in normal and STEMI because of elevated ST-segment and noise. In order to solve the above difficulty, we propose a way to effectively find the pulse using 12 leads. First, we apply the single QRS complex detector of Waveform Database (WFDB) library from Physionet [34], [35] with same parameter setting to 12-lead ECG respectively then apply it again after flipping the whole ECG upside down. An example of the procedure of QRS complex detection is shown in Fig. 7.

Following QRS detection, we summarize the locations of the QRS complexes and finally select the most likely location as QRS complex. In flipped ECG, Q-wave and S-wave can be regarded as R-wave by QRS complex detector. Due to above characteristic, our method can support to find the location of the QRS complex, even if the ST-segment is higher than QRS complex. Finally, we calculate the average interval $T$ between center points from selected QRS complex. Then, we segment the pulses by the center point of the QRS complex.
FIGURE 7. QRS detection block and detected QRS complexes.

FIGURE 8. Flow chart of the pulse segmentation. The QRS detection block is shown in Fig. 7.

FIGURE 9. Example of the QRS complex voting. The top subfigure shows the voting result. Middle and bottom of the figure show lead-I and lead-aVR view respectively. Selection criteria of QRS complex is most voted location. The selected QRS complexes are presented with red vertical dashes. We segment the pulse based on these criterion with average interval between red dashes.

with size $T$. The whole procedure of pulse segmentation is summarized in Fig. 8. An example of the QRS complex voting and selection is shown in Fig. 9.

C. CLASSIFICATION

We construct two models based on the two existing CNN architecture that already shows high performance in image classification. One of these architectures is VGGNet-16 from Simonyan and Zisserman [36] and the other is ResNet-34 from He et al. [37]. Although ResNet is already outperforming VGGNet using skip connection, we refer both architectures because we focus on improving performance rather than finding the better structure of the CNN. In Section IV, we experimentally show the performance improvement in both networks with our approach.

For handling the image data, the above two models are constructed with a 2-dimensional (2D) convolution layer. However, the 2D convolution is not appropriate for us. The color image has three channels of Red, Green, and Blue and each channel represents 2D (horizontal and vertical) signal. However, our data is 1-dimensional (1D) signal with 12 channels because ECG is basically 1-dimensional signal. When reconstructing the 12 channel ECG in 2D format with time axis as width and channel as height, 2D convolution can be used. But in that case, the correlations between channel are unnecessarily calculated. Thus, we use 1D convolution to construct CNN for focusing the information of each signal. The architecture of the modified CNN is shown in Fig. 10 and 11. The residual block for constructing 1D-ResNet-34 is shown in Fig. 12.

We present the training algorithm in Algorithm 1. The input data $x$ for CNN is segmented pulse from 12-lead ECG as shown in Fig. 4 and Fig. 5. and the target $y$ is the label corresponding to $x$. We construct the mini-batch $X$ and $Y$ with $x$ and $y$ for the training procedure. We define the loss function by cross entropy between output $\hat{Y}$ (mini-batch of $\hat{y}$) of CNN and target $Y$ as (1).

$$H(Y, \hat{Y}) = -\sum_{y \in Y, \hat{y} \in \hat{Y}} y \log \hat{y} \quad (1)$$

We minimize the loss as (1) using Adam optimizer [38]. We also initialize the parameters of the CNN using Xavier initializer prior to training [39].

Algorithm 1 Training Algorithm for CNN

Input and Target: $X$ (mini-batch of 12-lead data $X$) and $Y$ (mini-batch of label $y$ corresponding to $x$)

Output: $\hat{Y}$ (mini-batch of output $y$)

1: Initialize the neural network using Xavier initializer [39]
2: while the loss has not converged do
3:   Get $\hat{Y}$ by forward propagation in CNN
4:   Compute cross entropy between $Y$ and $\hat{Y}$
5:   Update parameters by Adam optimizer [38]
6: end while

IV. EXPERIMENTS

We compare how the performance varies by our approaches in this section. We deal with noise reduction and pulse segmentation sequentially. We also deal with the additional experiment to confirm diagnosis ability with a patient basis and how many data does CNN need for right performance. We provide the source code for the experiment at https://github.com/YeongHyeon/Enhancementing-Method-for-STEMI-Detection.

A. DATASET

The dataset is consisted of ECG records from SNUBH. The sampling rate of ECG is 500Hz and the total number of the
Prior to the experiment, we apply our preprocessing technique to all of the dataset. First, we call the set of the raw ECG as SNUBH-Raw (SNUBH-R) and we use this set as the baseline for comparing performance. We individually apply a notch filter and high-pass filter, and call these SNUBH-Notch (SNUBH-N) and SNUBH-Highpass (SNUBH-H) respectively. Lastly, we apply both the notch filter and high-pass filter and call this SNUBH-Both (SNUBH-B). All sets are consisted with 96 of normal and 179 of STEMI data.

Moreover, we segment pulses from SNUBH-R to SNUBH-B and call them SNUBH-R Pulse (SNUBH-RP) to SNUBH-B Pulse (SNUBH-BP). The pulse segmentation process is the same for all sets, but the number of pulses may differ because the location that regarded as QRS complex is not always the same depending on whether or not the noise reduction filter is applied. The amount of each dataset is shown in Table 2.

### B. NOISE REDUCTION VIEWPOINT

We conduct the experiment with SNUBH-R, SNUBH-N, SNUBH-H, and SNUBH-B in this section. We use Monte Carlo method to evaluate the STEMI detection performance [40]. We repeat the five-fold cross-validation [41] 30 times and calculate the mean and standard deviation of cross-validation to use them for a performance indicator. A higher mean indicates the better performance. By contrast, the lower the standard deviation is the better, and also stands for performance robustness. First, we measure the sensitivity and specificity shown in Table 3.

We show that the just applying noise reduction is not always effective to improve the performance. We measured the Area Under the Curve (AUC) of Receiver Operating Characteristic (ROC) curve [42] as an extra indicator in order to confirm that the noise reduction is effective. The measured AUC of ROC curve is summarized in Table 4.

We show that the just applying noise reduction is not always effective to improve the performance. We measured the Area Under the Curve (AUC) of Receiver Operating Characteristic (ROC) curve [42] as an extra indicator in order to confirm that the noise reduction is effective. The measured AUC of ROC curve is summarized in Table 4.

The AUC of ROC curve shown in Table 4 is not always increased although the noise reduction is applied as sensitivity and specificity. In addition, when the AUC is close to 1.0, the performance is generally considered to be excellent.
TABLE 4. Area under the curve (AUC) of receiver operating characteristic (ROC) curve with noise reduction. We present the AUC of ROC curve in the form of mean ± standard deviation.

<table>
<thead>
<tr>
<th>Set</th>
<th>Model</th>
<th>AUC of ROC curve</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNUBH-R</td>
<td>1D-VGGNet-16</td>
<td>0.586 ± 0.030</td>
</tr>
<tr>
<td></td>
<td>1D-ResNet-34</td>
<td>0.526 ± 0.029</td>
</tr>
<tr>
<td>SNUBH-N</td>
<td>1D-VGGNet-16</td>
<td>0.537 ± 0.025</td>
</tr>
<tr>
<td></td>
<td>1D-ResNet-34</td>
<td>0.537 ± 0.026</td>
</tr>
<tr>
<td>SNUBH-H</td>
<td>1D-VGGNet-16</td>
<td>0.635 ± 0.051</td>
</tr>
<tr>
<td></td>
<td>1D-ResNet-34</td>
<td>0.534 ± 0.030</td>
</tr>
<tr>
<td>SNUBH-B</td>
<td>1D-VGGNet-16</td>
<td>0.586 ± 0.036</td>
</tr>
<tr>
<td></td>
<td>1D-ResNet-34</td>
<td>0.516 ± 0.029</td>
</tr>
</tbody>
</table>

TABLE 5. Performance comparison with pulse segmentation after applying noise reduction. We present in the form of mean ± standard deviation.

<table>
<thead>
<tr>
<th>Set</th>
<th>Model</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNUBH-R</td>
<td>1D-VGGNet-16</td>
<td>0.842 ± 0.019</td>
<td>0.395 ± 0.053</td>
</tr>
<tr>
<td></td>
<td>1D-ResNet-34</td>
<td>0.832 ± 0.022</td>
<td>0.559 ± 0.037</td>
</tr>
<tr>
<td>SNUBH-N</td>
<td>1D-VGGNet-16</td>
<td>0.822 ± 0.014</td>
<td>0.559 ± 0.057</td>
</tr>
<tr>
<td></td>
<td>1D-ResNet-34</td>
<td>0.811 ± 0.014</td>
<td>0.485 ± 0.046</td>
</tr>
<tr>
<td>SNUBH-H</td>
<td>1D-VGGNet-16</td>
<td>0.930 ± 0.017</td>
<td>0.828 ± 0.075</td>
</tr>
<tr>
<td></td>
<td>1D-ResNet-34</td>
<td>0.928 ± 0.020</td>
<td>0.877 ± 0.023</td>
</tr>
<tr>
<td>SNUBH-B</td>
<td>1D-VGGNet-16</td>
<td>0.926 ± 0.012</td>
<td>0.867 ± 0.029</td>
</tr>
<tr>
<td></td>
<td>1D-ResNet-34</td>
<td>0.932 ± 0.011</td>
<td>0.896 ± 0.021</td>
</tr>
</tbody>
</table>

D. DIAGNOSIS WITH A PATIENT UNIT

We assess the diagnosing ability of our STEMI detection model on a patient basis. The diagnosing procedure is shown in Algorithm 2. We define the $X$ same as Algorithm 1 but in this situation, mini-batch $X$ is constructed with the pulses from one patient. The other variable $D$, $R$, $N$, $S$, and $\theta$ are defined as diagnosis of patient, a ratio of STEMI to total pulse, number of normal pulses, number of STEMI pulses, and threshold for diagnosis.

Algorithm 2 Diagnosing STEMI

Input: $X$ (12-lead pulses from one patient)

Output: $D$ (Diagnosis of patient)

1: Load the trained parameter to CNN
2: Classify the input $X$ to normal or STEMI using CNN
3: Count number of normal as $N$ and STEMI as $S$
4: Calculate $R$ by $S/(N+S)$
5: if $R$ $\geq$ $\theta$ then
6: $D$ = STEMI
7: else
8: $D$ = Normal
9: end if

We assess the performance with the same method as Section IV using the result of the above algorithm. Also, we consider improving the sensitivity by adjusting the threshold to reduce the patient missing rate in this experiment. The measured performance according to the threshold change is shown in Table 7.

TABLE 6. Area under the curve (AUC) of receiver operating characteristic (ROC) curve with pulse segmentation. We present the AUC of ROC curve in the form of mean ± standard deviation.

<table>
<thead>
<tr>
<th>Set</th>
<th>Model</th>
<th>AUC of ROC curve</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNUBH-R</td>
<td>1D-VGGNet-16</td>
<td>0.750 ± 0.034</td>
</tr>
<tr>
<td></td>
<td>1D-ResNet-34</td>
<td>0.738 ± 0.024</td>
</tr>
<tr>
<td>SNUBH-N</td>
<td>1D-VGGNet-16</td>
<td>0.716 ± 0.031</td>
</tr>
<tr>
<td></td>
<td>1D-ResNet-34</td>
<td>0.681 ± 0.025</td>
</tr>
<tr>
<td>SNUBH-H</td>
<td>1D-VGGNet-16</td>
<td>0.910 ± 0.024</td>
</tr>
<tr>
<td></td>
<td>1D-ResNet-34</td>
<td>0.936 ± 0.014</td>
</tr>
<tr>
<td>SNUBH-B</td>
<td>1D-VGGNet-16</td>
<td>0.923 ± 0.014</td>
</tr>
<tr>
<td></td>
<td>1D-ResNet-34</td>
<td>0.943 ± 0.010</td>
</tr>
</tbody>
</table>

We assess the performance with the same method as Section IV using the result of the above algorithm. Also, we consider improving the sensitivity by adjusting the threshold to reduce the patient missing rate in this experiment. The measured performance according to the threshold change is shown in Table 7.

TABLE 7. Performance of diagnosis on a patient basis in the form of mean ± standard deviation. The performance is measured with SNUBH-BP.

<table>
<thead>
<tr>
<th>Threshold</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.500</td>
<td>0.927 ± 0.013</td>
<td>0.904 ± 0.030</td>
</tr>
<tr>
<td>0.400</td>
<td>0.938 ± 0.013</td>
<td>0.899 ± 0.026</td>
</tr>
<tr>
<td>0.300</td>
<td>0.940 ± 0.012</td>
<td>0.892 ± 0.028</td>
</tr>
<tr>
<td>0.200</td>
<td>0.948 ± 0.010</td>
<td>0.881 ± 0.029</td>
</tr>
<tr>
<td>0.100</td>
<td>0.959 ± 0.009</td>
<td>0.828 ± 0.030</td>
</tr>
</tbody>
</table>

The sensitivity shows the increasing trend when the threshold is decreased, but the specificity is not. Thus, if physicians do not want to miss STEMI patients, we recommend dropping the threshold to get higher sensitivity instead of concession higher specificity.
E. TRAINING WITH FEW DATA

We check that the minimum amount of data to work with the proper performance. In fact, collecting and labeling the data requires a lot of cost. Moreover, medical data such as ECG is more difficult to collect because it can be collected only under certain circumstances.

Thus, if CNN performing well with only a few training dataset is more efficient. We fix the amount of test data with 10% of the total dataset assuming it is difficult to collect the data. Then we diversify the training data from 10% to 90% to find out how much data is required at least for the right performance. We use Monte Carlo method to evaluate but cross-validation is not used in this experiment. The performance is assessed after repeating the experiment 60 times as shown in Table 8.

As the ratio of training data to total data, i.e. the number of data used for training, increases, the sensitivity, specificity, and AUC of the ROC are improved simultaneously. The data used for training, increases, the sensitivity, specificity, in Table 8.

We check that the minimum amount of data to work with the proper performance. In fact, collecting and labeling the data requires a lot of cost. Moreover, medical data such as ECG is more difficult to collect because it can be collected only under certain circumstances.

Thus, if CNN performing well with only a few training dataset is more efficient. We fix the amount of test data with 10% of the total dataset assuming it is difficult to collect the data. Then we diversify the training data from 10% to 90% to find out how much data is required at least for the right performance. We use Monte Carlo method to evaluate but cross-validation is not used in this experiment. The performance is assessed after repeating the experiment 60 times as shown in Table 8.

As the ratio of training data to total data, i.e. the number of data used for training, increases, the sensitivity, specificity, and AUC of the ROC are improved simultaneously. The previous studies show their MI detecting performance. For example, Haraldsson et al. [8] and Strodthoff et al. [10] shows the sensitivity as 63.3% and 93.3%, and specificity as 85.0% and 89.7% respectively. When using 40% of dataset for training, the sensitivity and specificity outperform the previous studies that using 12-lead ECG. Thus, we conclude our algorithm appropriately works with only a small amount of data when obtaining a large amount of data is difficult.

F. PERFORMANCE IN PUBLIC DATA

Physionet provides various public ECG dataset [34]. In this section, we discuss with two datasets from Physionet that named as Long-Term ST (LTST) and Physikalisch-Technische Bundesanstalt Diagnostic ECG (PTB-ECG).

One of the related study use LTST dataset [6]. Basically, our purpose is detecting all localizations of STEMI that we already describe the reason for using 12-lead instead of selecting some leads in Section II-A. However, LTST dataset deal with ST-segment changes including ST-segment elevation and depression together. In addition, only two or three leads are contained in LTST dataset. Thus, we do not deal with LTST dataset to compare the performance.

For detecting MI, some of the studies use PTB-ECG dataset [4], [5], [10], [15], [17], [21]. The PTB-ECG dataset contains 12 leads and 3 Frank lead that calculated from standard leads system [43]. PTB-ECG dataset provides MI class without subtype distinction such as STEMI or not. There are STEMI within the MI class, and various other types of MI [1]. Therefore, the pulses of MI class in PTB-ECG dataset may be hard to be regarded as the same class. We confirm that different forms of ECG are existed in same MI class with sample data of PTB-ECG shown as Fig. 13.

Noise reduction and pulse segmentation is effective for our study because normal and STEMI are distinguished by elevated ST-segment in each pulse. However, in the case of PTB-ECG dataset, each data has different shape in the MI class assuming that the pulse segmentation is performed.

Performance comparison with equivalent approach is difficult because of the problem to be solved is different as described above. Thus, we experiment simply to confirm the effect of noise reduction approach with PTB-ECG dataset. We use the normal class (named as Healthy controls in PTB-ECG) and MI class only.

Unlike the dataset from SNUBH, each record of PTB-ECG dataset has different lengths, so we slice those to 5-second length. After slicing, we apply noise reduction and call these as PTB-Raw (PTB-R), PTB-Notch (PTB-N), PTB-Highpass (PTB-H), and PTB-Both (PTB-B) respectively. The frequency of power line in Korea is 60Hz, but in Germany, the source country of PTB-ECG, it is 50Hz. Therefore, we notch-filtered the harmonic frequency of 50Hz for PTB-ECG.

The amounts of ECG records for each dataset are 80 and 369 for normal and MI respectively. After noise reduction, the number of data for the normal class is 1880 and MI is 8207. We measured the performance with five-fold cross-validation using 1D-ResNet-34 with above datasets. We present result of performance measurement in Table 9.

We already achieved the cutting edge performance with raw data which named as PTB-R without preprocessing shown as Table 9. After applying notch filter the performance is slightly decreased. On the other hand, using a high-pass filter or using both filters improves performance. We confirm the high-pass filter is effective to enhancing the performance.
TABLE 9. Performance with PTB-ECG dataset and 1D-ResNet-34.
We present the mean performance of five-fold cross-validation.

<table>
<thead>
<tr>
<th>Set</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>AUC of ROC curve</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTB-R</td>
<td>0.999</td>
<td>0.928</td>
<td>0.987</td>
</tr>
<tr>
<td>PTB-N</td>
<td>0.999</td>
<td>0.916</td>
<td>0.982</td>
</tr>
<tr>
<td>PTB-H</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>PTB-B</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>

but notch filter is not. Because the raw data has very few mains hum. Like the experiment as Section IV-C, the notch filter is helpful to enhancing the performance.

V. CONCLUSION

In this paper, we propose an algorithm that potentially enables rapid, automate, and accurate diagnosis of STEMI with 12-lead ECG. Although we do not effort to find a better architecture of the CNN in this study, we suggest a simple method to improve the performance when using the same CNN architecture.

We use all leads of ECG to ease the limitations of previous studies. Further, we focus on enhancing the performance via noise reduction and pulse segmentation. Above preprocessing techniques are highly effective for enhancing the performance when using together. As a result, we assess our algorithm as high performance for diagnosing STEMI. Also, we expect the performance can be more improved when the optimal CNN architecture is found. In future studies, we consider improving our model to provide more detailed information such as localization of ST-Segment elevation.

APPENDIX A

DIFFERENCES IN PULSE SEGMENTATION

We already show the proposed method for QRS detection in Fig. 7 Section III-B, but we need to show the effectiveness of the method. Thus, We identify how our method helps in the pulse segmentation procedure. We apply the QRS complex detector of WFDB [35] to find peak-1 and peak-2.

We present the result of QRS detection and pulse segmentation sequentially. To confirm the usefulness of using peak-2, we use only peak-1 without peak-2 to vote the location of the QRS complex. We firstly show the result using peak-1 only in Fig. 14, then show the result using both peak-1 and peak-2 in Fig. 17. We also compare the above two methods for the same record in Fig. 20.

In a good case of Fig. 14, most QRS complexes are found. We exclude the first and last QRS complex because of fluctuated form by the high-pass filter. We expect that pulse segmentation can be performed correctly in the above situation. On the other hand, since QRS detection is not correctly performed in a bad case, pulse segmentation may not perform successfully because the QRS complex is not selected correctly. The reason for failure is that the other waves are elevated than R-wave in the most channel of ECG as shown in Fig. 15 and Fig. 16.

Also, Fig. 17 shows the same context as Fig. 14. Even if a specific wave has an almost similar form as the R-wave, QRS detection is disturbed. Also, QRS detection is hindered by the frequent appearance of a towering form, for example, the P-wave in the third in Fig. 17.

The QRS detection performance can be different in the same record by the method. Our method uses the peak-2 by flipping upside down the ECG for voting the location of QRS complex. We compare the results with and without peak-2 respectively as follows.

The pulse segmentation results are shown in Fig. 18 and Fig. 19. The bad cases come from the failure of QRS detection, same as Fig. 15 and Fig. 16.

In order to facilitate comparison, we use the same record Pat-F for comparison. The QRS detection results are different for two method. The first figure of Fig. 20, since the first method uses only peak-1 at voting, T-wave can be considered

FIGURE 14. The lead-I view of QRS detection result using only peak-1 for voting. From top to bottom, we name each record as Patient-A (Pat-A), B, C, and D. Upper two figures, show the good case of QRS detection for normal and STEMI class respectively. Lower two figures show bad cases for each class. We make an exception the first and the last QRS complex because it has the fluctuated form caused by a high-pass filter.
as R-wave. Thus, most of the selected location is pointing T-wave.

In the lower case of the Fig. 20, the QRS complexes are successfully found. Because the final locations for considering as QRS complex are correcting by S-wave. We present the pulse segmentation result from Fig. 20 as shown in Fig. 21. We confirm the effectiveness of using peak-2 in this comparison procedure.

When using peak-1 only, the accuracy is 0.999 for normal class and 0.824 for STEMI class. The accuracy for normal class is slightly decreased as 0.997 but in STEMI class the accuracy is highly increased as 0.983. The average accuracies are 0.912 for using only peak-1 and 0.990 for using both peak-1 and peak-2. The accuracy of QRS segmentation is improved by using peak-2 with peak-1.

The high accuracy means that the data of each class are highly consistent. Classification performance of CNN is depending on the data consistency of each class. The higher consistency helps more improve the performance, so our method, using both peak-1 and 2, is effective for performance enhancement.

APPENDIX B
PULSE SEGMENTATION WITH MISSING LEAD ECG
The ECG dataset from the SNUBH is recorded in real-world clinical practice, including the emergency room and...
FIGURE 21. The pulse segmentation example from Fig. 20. These pulses are segmented from the same record Pat-F but show the different result.

TABLE 10. The number of segmented pulse in lead missing situation. We calculate the average number of pulses by conducting the Monte Carlo method 30 times. The average amount of the pulse for each class and each situation is almost same as the SNUBH-BP.

<table>
<thead>
<tr>
<th>Number of removed lead</th>
<th>Normal</th>
<th>STEMI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non removed (SNUBH-BP)</td>
<td>899</td>
<td>1805</td>
</tr>
<tr>
<td>1 of 12</td>
<td>906</td>
<td>1805</td>
</tr>
<tr>
<td>2 of 12</td>
<td>902</td>
<td>1809</td>
</tr>
<tr>
<td>3 of 12</td>
<td>903</td>
<td>1798</td>
</tr>
<tr>
<td>4 of 12</td>
<td>902</td>
<td>1797</td>
</tr>
<tr>
<td>5 of 12</td>
<td>901</td>
<td>1786</td>
</tr>
<tr>
<td>6 of 12</td>
<td>899</td>
<td>1774</td>
</tr>
<tr>
<td>7 of 12</td>
<td>900</td>
<td>1767</td>
</tr>
<tr>
<td>8 of 12</td>
<td>894</td>
<td>1737</td>
</tr>
<tr>
<td>9 of 12</td>
<td>894</td>
<td>1737</td>
</tr>
<tr>
<td>10 of 12</td>
<td>885</td>
<td>1672</td>
</tr>
<tr>
<td>11 of 12</td>
<td>983</td>
<td>1861</td>
</tr>
</tbody>
</table>

outpatient clinic. As a rule, if one or more leads are missing, the ECG should be re-recorded. Thus, the imperfectly recorded ECG will not be used for diagnosing STEMI or not in the clinic.

However, we conduct the experiment for confirming that our pulse segmentation algorithm works properly whether some of the leads are missed or not. We create a situation of missing up to three leads in the 12-lead ECG. Each missing situation is generated randomly.

We firstly, check the number of pulses in each situation where some leads are removed. The result of simulation is shown in Table 10. As shown in Table 10, the number of the pulse is usually similar to the baseline (SNUBH-BP). The pulse segmentation accuracy of the SNUBH-BP is 0.990 that described in Appendix A. Thus, when the number of segmented pulses for each situation is similar to the number of pulses of the SNUBH-BP set as a reference, it can be regarded the pulse segmentation is performed well.

We show the sample pulse segmentation results when three leads are randomly removed in Fig. 22. As shown in the figure, pulse segmentation is properly performed whether some lead does not appear. The pulses are properly segmented even if other leads are randomly removed in the same way.

The pulse segmentation method performs well even if some leads are partially removed. Because we utilize 12 leads and voting the location of the QRS complex. In the real world, imperfectly recorded ECG is not used to diagnosing, but our algorithm can find and segment the pulse for all leads. In conclusion, it can be said that our pulse segmentation algorithm is effective when the purpose is only finding QRS complex or pulse segmentation.

APPENDIX C
COMPARISON WITH OTHER CLASSIFIER

In Section II-C, we already describe the reason for using CNN-based algorithm. However, we conduct some extra experiment using conventional classifiers as kNN, SVM, and DT for showing the benefit of our preprocessing method. Also, this experiment can show the limitation of the above models mentioned before. We conduct 30 times experiments with the Monte Carlo method and the result is shown in Fig. 23.

For the experiment, the k (number of nearest neighbors) of kNN and the maximum depth (maximum number of pruning) of DT are adjusted. The maximum sensitivity, specificity, and AUC of the ROC curve are 0.853, 1.000, and 0.894. However, 1.000 of specificity is achieved by abandon the sensitivity. The sensitivity and specificity are measured by a static threshold of 0.5 so the performance may differ when the threshold is changed. Thus, we focus on the AUC that can show the summarized performance. Because AUC is measured by the various threshold.

Fig. 23 shows that the preprocessing method is also effective for other classifiers by an increasing trend of AUC. The performance is still lack than CNN-based classifier. For better performance, additional efforts have been required,
FIGURE 23. The measured performance for confirming the effectiveness of noise reduction and pulse segmentation. Experiments are conducted with 8 kinds of the dataset (SNUBH-R to SNUBH-BP) and 3 kinds of classifier (kNN, SVM, and DT) for comparing the performance of each classifier. For describing the dataset, SNUBH-R, N, H, and B represent raw, notch, high-pass, and both filters. SNUBH-RP, NP, HP, and BP are the pulse segmented datasets from R, N, H, and B, respectively. Each figure shows sensitivity, specificity, and AUC of ROC curve for each classifier sequentially. Each box plot presents maximum, first quartile (Q1), mean, third quartile (Q3), and minimum performance sequentially. The performance indicators show the increasing trend by adapting our preprocessing method.

for example, find k for using kNN and a maximum depth of DT. When using SVM, the above cost may ease but SVM also needs to find better features extraction method for distinguishing each class. However, we achieve much better performance by segmenting core information and using a CNN-based algorithm. That means CNN-based algorithm reduces most of the cost for find or construct the better classifier.
REFERENCES


YEONGHYEON PARK received the B.S. degree from the Department of Digital Information Engineering, Hankuk University of Foreign Studies, Yongin, South Korea, in 2018. He is currently pursuing the M.S. degree with the Department of Computer and Electronic Systems Engineering, Hankuk University of Foreign Studies, Yongin, South Korea. His research interests include medical data analysis and artificial intelligence.

IL DONG YUN (S’89–M’97) received the B.S., M.S., and Ph.D. degrees in control and instrumentation engineering from Seoul National University, Seoul, South Korea, in 1989, 1991, and 1997, respectively. From 1996 to 1997, he was with Daewoo Electronics, Seoul, South Korea, as a Senior Engineer. Since 1997, he has been with the Hankuk University of Foreign Studies, Yongin, South Korea. His research interests include medical image analysis and computer vision.

SI-HYUCK KANG received the B.S. and M.S. degrees with the Department of Medicine, Seoul National University, Seoul, South Korea, in 2005 and 2013, respectively. He is currently pursuing the Ph.D. degree with the Department of Medicine, Seoul National University, Seoul, South Korea. Since 2013, he has been with the Cardiovascular Center, Internal Medicine, Seoul National University Bundang Hospital, Seongnam, South Korea.

***