Synthesis of Electrocardiogram V-Lead Signals from Limb-Lead Measurement using R-peak Aligned Generative Adversarial Network

JeeEun Lee[1], KyeongTaek Oh[2], Byeongnam Kim[3] and Sun K. Yoo[*]

***Abstract*—Recently, portable electrocardiogram (ECG) hardware devices have been developed using limb-lead measurements. However, portable ECGs provide insufficient ECG information because of limitations in the number of leads and measurement positions. Therefore, in this study, V-lead ECG signals were synthesized from limb leads using an R-peak aligned generative adversarial network (GAN). The data used the Physikalisch-Technische Bundesanstalt (PTB) dataset provided by PhysioNet. First, R-peak alignment was performed to maintain the physiological information of the ECG. Second, time domain ECG was converted to bi-dimensional space by ordered time-sequence embedding. Finally, the GAN was learned through the pairs between the modified limb II (MLII) lead and each chest (V) lead. The result showed that the mean structural similarity index (SSIM) was 0.92, and the mean error rate of the percent mean square difference (PRD) of the chest leads was 7.21%.**

***Index Terms*— Electrocardiogram, Generative adversarial network, Synthesis**

I. Introduction

Electrocardiograms (ECGs) are applied to various situations ranging from heart rate monitoring in daily living, to patient monitoring in operation rooms, and to the diagnosis of arrhythmia [1]. Recently, various portable ECG hardware devices that allow easy measurement have been developed on the basis of the limb-lead ECG method. These devices enable the users to perform ECG measurement and monitoring at home. The portable ECG devices are applicable to the mobile healthcare environment, because they are not limited by time and space [2].

The number of leads and measurement positions are dependent on the ECG application. The representative ECG measurement methods include the standard limb-lead method and the chest-lead method. In the standard limb-lead method, leads are attached to the right arm, left arm, and left leg to perform the measurement by bipolar derivation. In the chest-lead method, chest(V) leads should be attached to parts near the heart at accurate positions. Beside these methods, ECG may be performed using the standard 12-lead method, simple three-electrode monitoring, five-electrode monitoring, Lewis-lead method, and Fontaine-lead method [1]. V leads, which are for the measurement of myocardial motion, are used to examine cardiac diseases including myocardial infarction [3]. However, because the chest-lead method requires that many leads be attached, the measurement is inconvenient; thus its application for a wearable ECG device is limited. Therefore, portable ECG may not provide sufficient ECG information due to the limitation in the number of leads attached. This limits wearable ECG to usage only at homes, and makes it difficult to analyze the ECG data clinically [2]. Therefore, the present study was conducted to convert a representative limb signal of the modified limb Ⅱ (MLⅡ) lead to mimic the chest-lead ECG signals obtained from many leads.

Various studies have been conducted for ECG-lead conversion to overcome the relevant problems. With the Dower transformation and affine transformation, coefficients were calculated by the least-square method to perform the lead data reconstruction [2], [3], [4], [5]. Principal component analysis (PCA) allows reconstruction of signals through characteristics vectors that project the principal components [5]. However, these methods involve complicated calculations and are limited in sensitivity depending on individual differences and measurement positions.

Various deep learning methods have been applied to the ECG synthesis, including deep neural networks (DNNs), convolution neural networks (CNNs), recurrent neural networks (RNNs), and generative adversarial networks (GANs). These methods employ non-linear transformation and thus are appropriate for data having diversity resulting from variation in individuals and measurement positions. However, these methods are currently applied to noise recovery, arrhythmia classification and ECG signal prediction, and are rarely used in the synthesis of multi-lead ECG data [6], [7], [8], [9].

Recently, GAN has shown excellent performance in image style translation [10]. While for the conventional CNN, learning the Euclidean distance is a blur; a GAN used to perform loss-based learning, gives clearer results than those obtained using CNN [11], [12]. However, because the GAN is currently optimized for images, its application to time domain signals is difficult. Therefore, the ECG signal processing is critical to the application of GAN to ECG signals, which are time domain signals. In the present study, ECG signals were

[1] Graduate Program of Biomedical Engineering, Yonsei University, Seoul, Korea (e-mail: jeunlee@yuhs.ac).

[2] Graduate Program of Biomedical Engineering, Yonsei University, Seoul, Korea (e-mail: okt2704@yuhs.ac)

[3] Graduate Program of Biomedical Engineering, Yonsei University, Seoul, Korea (e-mail: bnkim007@yuhs.ac)

* Department of Medical Engineering, Yonsei University College of Medicine, Seoul, Korea (e-mail: SUNKYOO@yuhs.ac).

***** Correspondence: sunkyoo@yuhs.ac; Tel.: +82-2-2228-1919

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2019.2936583, IEEE Journal of Biomedical and Health Informatics

converted into a bi-dimensional space. In particular, R-peak alignment and ordered time-sequence embedding were applied to maintain the semantic information of the ECG signals. In addition, the GAN was made to undergo learning through ECG lead-pairs to synthesize the chest-lead data.

II. METHOD

A. Datasets

The data used in the present study were the data of 52 healthy individuals obtained from the Physikalisch-Technische Bundesanstalt (PTB) dataset provided by PhysioNet. The PTB dataset is an ECG dataset of less than two minutes, including both the standard limb-lead and chest-lead ECG data [13]. The data analyzed in the present study were the ECG signals obtained from the MLII leads and V leads. The sampling rate of the original data was 1000 Hz, but it was down-sampled to 250 Hz. Of the total, 70% of the data was used for the training set, 20% for the validation set, and 10% for the testing set.

B. R-peak Alignment

ECG signals are periodic and thus vary according to the heart beat patterns. Therefore, semantic information may be extracted by analyzing the ECG patterns. ECG signals of 1 s have the P-Q-R-S-T information, which is the key pattern of a heart beat [14]. The width of QRS duration, and QT interval are important indicators of ECG analysis. Since the heart rate is 70 / min on average, the pattern of consecutive P-Q-R-S-T should be synthesized for feature extraction. Therefore, the leads used in the present study were segmented by synchronizing the ECG pattern before and after 500 ms with reference to the R-wave, which is the median value of P-Q-R-S-T pattern of the MLII lead. R-peaks were extracted based on the Pan-Tompkins algorithm. After squaring the differentiation values of the bandpass filter signal, it performs moving window integration to apply the adaptive threshold to the QRS amplitude and the RR interval [15]. The segmented signal based on the R-peak is expressed in Equation (1) where x_p represents the sample at the time when the MLII signal has the R-peak. Here, $\mathbb{E}_i$ denotes one epoch ECG after being segmented (with 1 s of ECG information) and a total of 250 samples. N is the number of epochs, and it depends on the time that the ECG is measured. Therefore, $\mathbb{E}_i$ retains the semantic information through synchronization of the R-peak alignment-segmented ECG signals.

$$\mathbb{E}_{i=1,\cdots,N} = \{x_{p-125}, x_{p-124}, \cdots, x_{p-1}, x_p, x_{p+1}, \cdots, x_{p+123}, x_{p+124}\} \quad (1)$$

C. Augmentation

The application of the PTB dataset to a GAN requires two types of augmentation.

The first is augmentation to convert the time domain ECG signals, used as the input to the GAN, into a bi-dimensional space. In the present study, the time domain ECG signals were converted to a bi-dimensional space having the size 256×256. The x-axis of the bi-dimensional space represents the ECG signals of the R-peak aligned epochs, and the y-axis represents the ECG signals processed by ordered time sequence embedding, which show the sequential ECG flow information of an individual. Therefore, both the x-axis and y-axis represent time-dependent information.

The PTB dataset, which contains short signals of less than two minutes, gives about 120 epochs of segmented ECG signals for an individual [13]. Maintaining the sequential information is important because the ECG signals are time domain signals. Therefore, the data augmentation was performed while retaining the time-dependent information of the segmented ECG signals through ordered time-sequence embedding. Ordered time-sequence embedding is used to convert one epoch of ECG signals to a bi-dimensional space through linear interpolation according to the time sequence. The ECG signals are augmented from 120 epochs to 256 epochs. To include time-dependent information as much as possible, the interpolation coefficient, the smallest unit, is fixed at three, and each epoch is copied to provide a total of three copies for linear interpolation. Therefore, for example, 120 epochs contained in an ECG signal are augmented to 360 epochs. In addition, to convert the signals to a 256x256 bi-dimensional space, $\mathbb{E}_i$ having 250 samples is padded with zero and then extended to 256 samples. This process is shown in Equation (2) where $\mathbb{S}$ denotes the ECG signals of the healthy control data of one individual after being ordered time-sequence embedded. The data size is $3\mathrm{N} \times 256$, where N is the number of epochs. Therefore, if the interpolation coefficient is three, the conversion to a 256x256 bi-dimensional space requires data for at least 1 min and 26 s.

$$\mathbb{S} = \mathrm{concat}\left(\begin{bmatrix} \mathbb{E}_{x,y=1} \\ \vdots \\ \mathbb{E}_{x,y=3\times N} \end{bmatrix}\begin{bmatrix} 0_{x=251,\cdots 256,y=1} \\ \vdots \\ 0_{x=251,\cdots 256,y=3\times N} \end{bmatrix}\right) \quad (2)$$

The size of the PTB dataset used for the training of the GAN model was insufficient. Therefore, to augment the number of samples, each of the interpolated epochs was moved to extract the samples converted into a bi-dimensional space. Therefore, if the original ECG data had 120 epochs, the data were interpolated to 360 epochs, and 105 samples could be extracted depending on the epoch moving. This process is expressed in Equation (3) where N denotes the number of epochs contained in the data of each individual.

$$Number\ of\ Sample = 3 \times N - 256 + 1 \quad (3)$$

Through these processes, the signals obtained from the MLII lead and the V leads were converted into a bi-dimensional space. Each of the converted V-lead samples was paired with the MLII-lead sample to be used as the input data of the GAN.

D. GAN Architecture

The GAN consisted of a generator synthesizing the V leads as well as a discriminator that distinguished a true signal from a

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2019.2936583, IEEE Journal of Biomedical and Health Informatics

false signal, and underwent learning through the pair dataset. The learning by the GAN was performed using the reconstruction loss, which generates signals that are similar to the ground truth, and the adversarial loss, which seems like the truth. The architecture and hyper parameters of the GAN used in this study are based on the structure of the image-to-image translation GAN [12]. In particular, the generator was changed using a nine-block residual network to retain information consistently. The residual network has two 3X3 filter sizes and convolution layers with reflect padding. After the convolution, it performs a batch normalization process. Also, a rectifier linear unit (ReLU) is used as the activation function of the first convolution layer [16]. The discriminator recovered the detailed information depending on the R-peak-aligned ECG pattern by applying the patch GAN [12].

Equations (4) to (6) were used to obtain the loss of the GAN. R_MLII represents the MLII-lead sample, R_v the V-leads sample, and G_v the regenerated V-leads sample. Equation (6) was used to calculate the loss used in the present study. The loss was calculated by adding the error between the actual V-leads sample and the absolute value of the generated V-leads sample to the GAN loss calculated in Equation (5). Notably, the error was minimized by calculating the reconstruction loss depending on the number of pixels of the data pairs through Equation (4), and thus the performance of data recovery is better with the GAN model developed in the present study, than with the conventional GAN model [12].

$$L_1(G) = \mathbb{E}[\|R_v - G_v\|_1] \quad (4)$$

$$L_{vGAN}(G, D) = \mathbb{E}[logD(R_{MLII}, R_v)] + \mathbb{E}[\log(1 - D(R_{MLII}, G_v)] \quad (5)$$

$$G^* = arg \min_G \max_D L_{vGAN}(G, D) + \lambda L_1 \quad (6)$$

Figure 1 shows the overall structure of the R-peak-aligned GAN used in the present study. For the R-peak aligned GAN training, we set the batch size to 1 and used the minibatch stochastic gradient descent (SGD). The learning rate was set to 0.0002 and the hyper parameter optimization was performed through adaptive moment estimation (Adam).

E. Evaluation

1) Evaluation Datasets

In the present study, R-peak alignment, ordered time sequence embedding, and paired dataset were used to recover the V-leads data from the MLII lead. The R-peak alignment retains the semantic information of the ECG pattern. The necessity for the R-peak alignment was verified by comparing the data of those that had, and had not, undergone the R-peak synchronization. In addition, the effect of ordered time-sequence embedding, which was performed to reflect the long-term information of the ECG signals, was verified by comparing the data that had undergone random embedding with those that were entirely embedded using a single epoch. In addition, to examine the superiority of using the paired dataset, the results were compared with those obtained using the cycleGAN model to perform the generation from datasets without pairs [17].

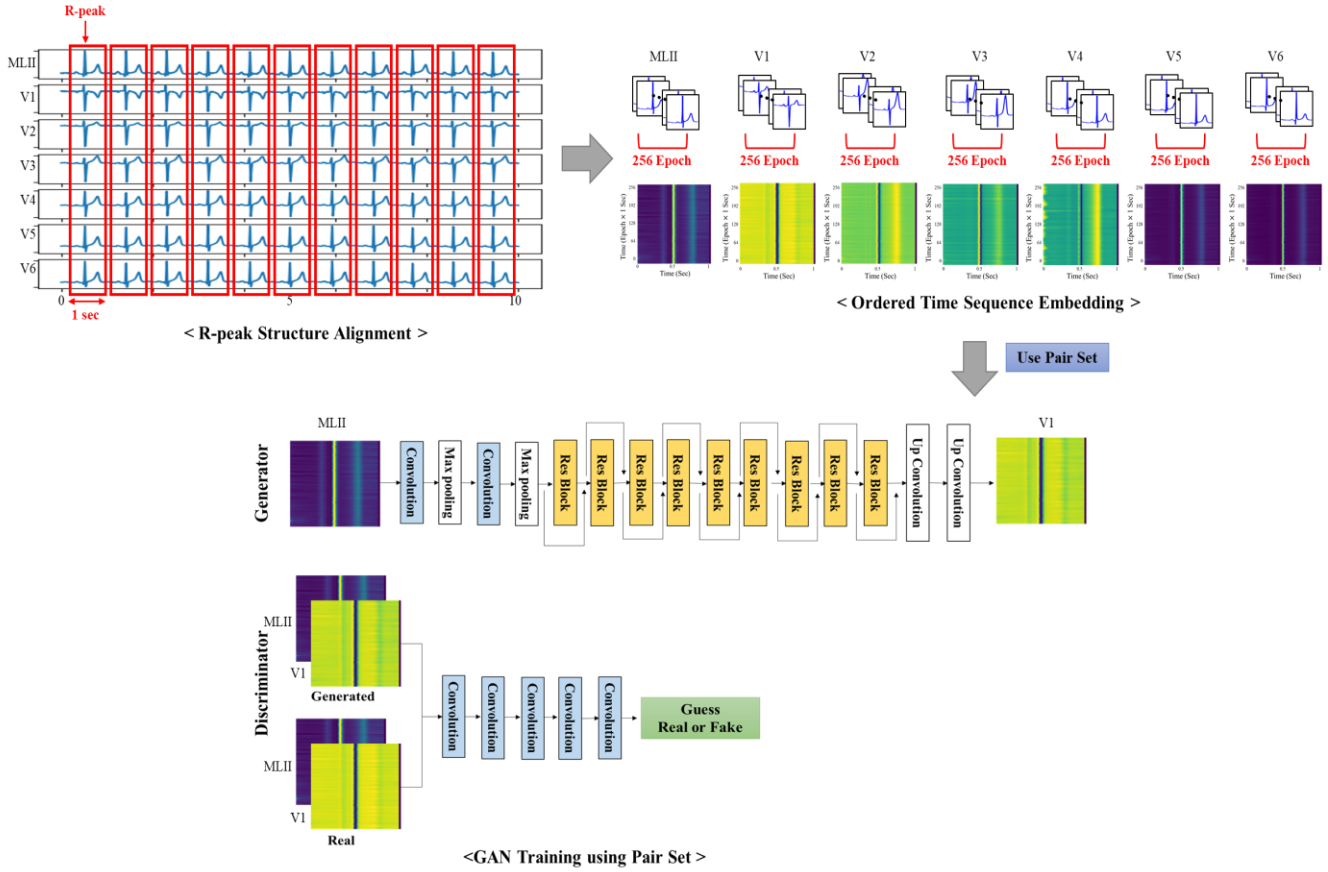

Fig. 1. Overall Architecture of an R-peak Aligned GAN for Synthesis of ECG V Leads

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2019.2936583, IEEE Journal of Biomedical and Health Informatics

The ECG signals processed by the method developed in the present study were labeled R-peak-alignment time embedding (RATE). Those that underwent R-peak alignment were labeled R-peak- nonalignment time embedding (RNTE). Those that underwent R-peak alignment and included random long-term information were labeled R-peak-alignment random time embedding (RARE). Those that underwent time-dependent alignment and were entirely embedded using a single epoch were labeled R-peak-alignment one-epoch embedding (RAOE) and those that were obtained by applying the cycleGAN were labeled R-peak-alignment time embedding with cycleGAN (RATEC). Therefore, the V-leads ECG signals recovered using these five methods were compared.

2) Subjective Evaluation Metrics

Subjective evaluation metrics (SEM) were measured in the present study to compare the results of the ECG recovery. The SEM were used to evaluate a GAN model about how visually similar a generated image was to the actual image.

The inception score (IS), obtained by inputting a generated image to an inception net, was used to evaluate an image with reference to its quality and diversity. A higher IS means successful recovery of the ECG data (a value around two) [18]. The IS value was used to measure the performance in generating an ECG image.

When only one image is generated (depending on the class), the IS may misrepresent the data and thus has the risk of overfitting. Therefore, the structural similarity index (SSIM) was used to measure the structural similarity between images against image distortion. The SSIM is the product of luminance, contrast, and structure. An SSIM value closer to one means that the generated image is more similar to the original one, while a value closer to zero means that the generated image is more different from the original one [19]. Because the ECG signals contain semantic information, it is important to retain their structure. Therefore, the performance of a generated ECG image in retaining its structure is measured with reference to the SSIM value.

3) Qualitative Evaluation Metrics

The qualitative evaluation metrics (QEM) show the error between the recovered ECG signals and the original signals. The QEM method was used to investigate the correlation between the actual one-dimensional ECG signals extracted from the image and the reconstructed one-dimensional ECG signals. The indices employed were percent mean square difference (PRD), correlation coefficient (Corr), and amplitude difference coefficient (Amp). The PRD represents the error between the original ECG signals and the generated ECG signals, a smaller PRD value meaning more successful recovery [20]. The Corr value represents the similarity between the original ECG signals and the generated ECG signals, and the Amp value shows the amplitude difference between the two signals [2]. The PRD, Corr, and Amp are calculated using Equations (7) to (9), respectively, where N is the number of samples, R_v the original V-leads signals, and G_v the V-leads signals generated by the generator [2], [20].

$$\mathrm{PRD}(\%) = 100 \times \sqrt{\frac{\sum_{n=1}^{N}(R_v(n)-G_v(n))^2}{\sum_{n=1}^{N}(R_v(n))^2}} \qquad (7)$$

$$\mathrm{Corr} = \left\{\frac{\sum_{n=1}^{N} R_v(n)\times G_v(n)}{\left(\sum_{n=1}^{N}(R_v(n))^2 \times \sum_{n=1}^{N}(G_v(n))^2\right)^{1/2}}\right\} \qquad (8)$$

$$\mathrm{Amp} = \left\{\frac{\sum_{n=1}^{N} R_v(n)\times G_v(n)}{\sum_{n=1}^{N}(R_v(n))^2}\right\} \qquad (9)$$

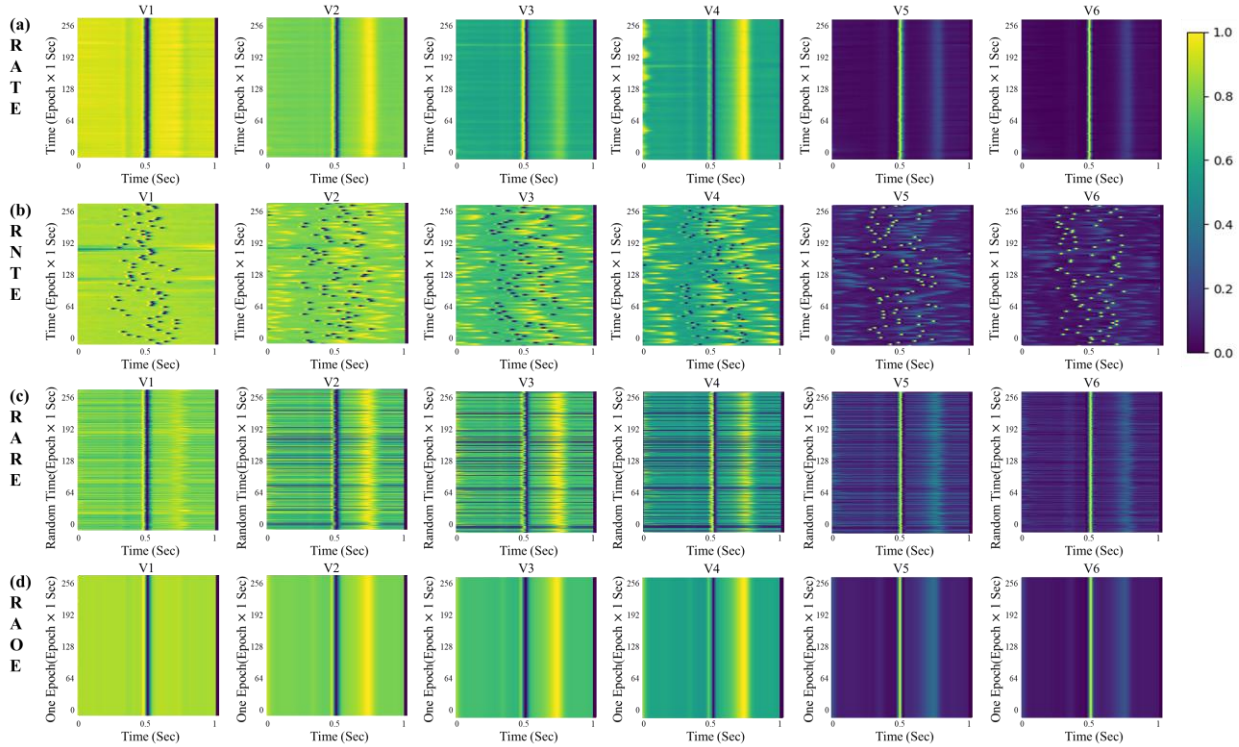

Fig. 2. Bi-dimensional Space Transformation Results: (a) RATE, (b) RNTE, (c) RARE, (d) RAOE

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2019.2936583, IEEE Journal of Biomedical and Health Informatics

III. Results

A. Bi-dimensional Space Transformation

Figure 2 shows the result of the conversion of the V-leads signals into a bi-dimensional space. Figure 2(a) shows the RATE result, Figure 2(b) the RNTE result, Figure 2(c) the RARE result, and Figure 2(d) the RAOE result. As shown in Figure 2, different images were generated depending on the application of R-peak alignment and ordered time-sequence embedding. The ECG signals were converted into images

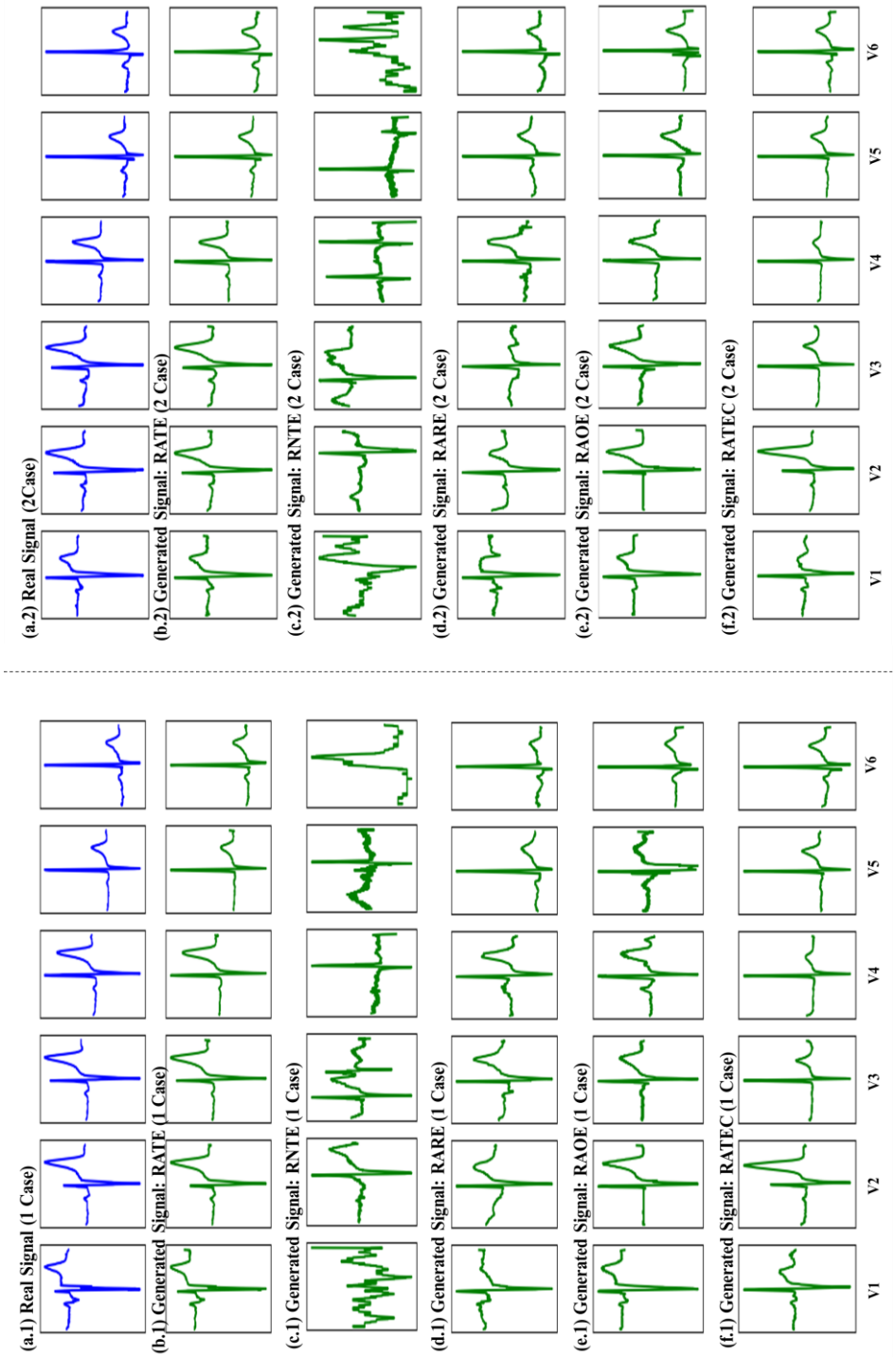

Fig. 3. Two Cases of Generated ECG: (a) Real V ECG, (b) Generated ECG from RATE, (c) Generated ECG from RNTE, (d) Generated ECG from RARE, (e) Generated ECG from RAOE, (f) Generated ECG from RATEC

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2019.2936583, IEEE Journal of Biomedical and Health Informatics

having the pixel size 256x256. Each image was expressed by pseudo-coloring of the pixels with a value between "0" and "1". Each of the V-leads ECG signals was paired with the imaged obtained from the MLII lead.

B. Generated ECG

Figure 3 shows the generated V-leads ECG signals from two data cases where the actually measured signals were colored blue and the synthesized signals were colored green. Figure 3(a) shows the original V-leads ECG signals, Figure 3(b) the RATE signals, Figure 3(c) the RNTE signals, Figure 3(d) the RARE signals, Figure 3(e) the RAOE signals, and Figure 3(f) the lead signals recovered through the RATEC. Visual inspection of the regenerated signals showed that the RATE signals were not significantly different from the original signals. The RNTE signals were the most different from the original signals. As shown in V6 of Figure 3(c), the RNTE signals even became angulated. In the RARE, RAOE, and RATEC signals, the patterns of the P-Q-R-S-T wave, which have large amplitude, were recovered, but the patterns of the original signals were not perfectly recovered. In particular, as shown in V6 of Figure 3(d.1), (d.2) and Figure 3(e.1), (e.2), the peak amplitude of the Q-wave was swapped with that of the S-wave. Also, V5 of Figure 3(e.1) generated a wrong signal, but V5 of Figure 3 (e.2) was synthesized a similar signal compare to the original signal. The generated RATEC signals looked like well-recovered V-leads signals; however, because learning was performed with respect to the dataset, only general V-leads signals were generated. Comparison of these with the real signals indicates that the recovery was flawed.

C. Result of Subjective Evaluation Metrics

Table 1 shows the mean IS and SSIM values of the generated ECG images.

The signals having the highest IS value in each lead are written in bold type. Except for the V1 lead, the highest IS and standard deviation value are found in the RATE method for all leads. For the V1 lead, the RNTE result shows the greatest IS value, which is about 0.76 greater than that of the RATE result. The RNTE result shows the second highest IS values for the V2, V3, V4, and V5 leads. In contrast, the RARE result shows the lowest IS and standard deviation values for all leads.

In Table II, the highest SSIM values are also written in bold type. The SSIM values after the final calculation are highest in the RATE signals for all leads, except for the V6 lead. The SSIM value of the RATE signals is higher than 0.9 for all leads. The SSIM value of the RAOE signals for the V6 lead is higher by 0.3 than that of the RATE signals. The mean SSIM values of the RAOE signals are higher than those of other signals, except for the RATE signals. In contrast, the SSIM of the RNTE signals is lower than that of other signals for all leads, and this is lowest for the V6 lead (0.07).

TABLE I
RESULT OF SUBJECTIVE EVALUATION METRICS

		V1	V2	V3	V4	V5	V6	Mean
IS	RATE	1.70	**1.71**	**1.72**	**1.82**	**1.64**	**1.57**	**1.69**
	RNTE	**1.76**	1.53	1.62	1.49	1.44	1.27	1.52
	RARE	1.11	1.08	1.08	1.12	1.09	1.11	1.10
	RAOE	1.22	1.26	1.30	1.37	1.19	1.28	1.27
	RATEC	1.12	1.15	1.30	1.23	1.10	1.11	1.17
SSIM	RATE	**0.92**	**0.93**	**0.90**	**0.90**	**0.94**	0.90	**0.92**
	RNTE	0.25	0.23	0.18	0.21	0.33	0.07	0.21
	RARE	0.37	0.21	0.14	0.17	0.27	0.50	0.28
	RAOE	0.91	0.87	0.86	0.89	0.81	**0.93**	0.88
	RATEC	0.58	0.65	0.72	0.65	0.79	0.78	0.70

TABLE II
STATISTICAL EVALUATION OF SUBJECTIVE EVALUATION METRICS USING ONE WAY ANOVA

	IS				SSIM			
	df	MS	F	p-value	df	MS	F	p-value
	RNTE	RARE	RAOE	RATEC	RNTE	RARE	RAOE	RATEC
V1	4	0.234	23.393	<0.0001	4	81.613	18874.207	<0.0001
	p<0.05	p<0.05	p<0.05	p<0.05	p<0.05	p<0.05	p<0.05	p<0.05
V2	4	0.088	7.325	<0.0001	4	104.092	29874.249	<0.0001
	p>0.05	p<0.05	**p>0.05**	**p>0.05**	p<0.05	p<0.05	p<0.05	p<0.05
V3	4	0.113	11.197	<0.0001	4	123.103	34788.095	<0.0001
	p>0.05	p<0.05	p<0.05	**p>0.05**	p<0.05	p<0.05	p<0.05	p<0.05
V4	4	0.119	7.373	<0.0001	4	112.722	29509.326	<0.0001
	p>0.05	p<0.05	p<0.05	**p>0.05**	p<0.05	p<0.05	p<0.05	p<0.05
V5	4	0.085	24.755	<0.0001	4	82.270	15447.012	<0.0001
	p<0.05	p<0.05	**p>0.05**	**p>0.05**	p<0.05	p<0.05	p<0.05	p<0.05
V6	4	0.042	9.651	<0.0001	4	114.190	45666.400	<0.0001
	p<0.05	p<0.05	**p>0.05**	p<0.05	p<0.05	p<0.05	p<0.05	p<0.05

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2019.2936583, IEEE Journal of Biomedical and Health Informatics

Table 2 shows statistical results for IS and SSIM values. One-way ANOVA test was performed to confirm the significance of each RATE, RNTE, RARE, RAOE and RATEC scheme for each of the generated V leads. The IS and SSIM values calculated from the samples generated by each method were used as inputs to the one-way ANOVA test. Since the number of n groups in each group was the same, Bonferroni post hoc test was used to confirm that RATE was significantly different from other methods (RNTE, RARE, RAOE, RATEC). Both IS and SSIM were confirmed to be significant by one-way ANOVA test ($p < 0.0001$). However, post hoc test results showed that RATE was significantly different from RARE ($p < 0.05$), but not significantly different from RNTE, RAOE, and RATEC depending on lead. Especially, RATEC was not significant in the case of lead (V2, V3, V4, V5). On the contrary, in the SSIM index, post hoc test results show that RATE is significantly different from all other results ($p < 0.05$).

D. Result of Qualitative Evaluation Metrics

Table 3 shows the QEM result of the ECG signals generated by the R-peak structure alignment. The smallest PSD values, representing error between the actual signals and the generated signals (the Corr and Amp values closest to "1"), are written in bold type. The mean Amp value is the mean of the absolute values of the difference between "1" and the Amp values.

The mean PSD of the RATE signals was 7.2%, which is the smallest. In particular, the RATE signals show the smallest error (4.32%) for the V1 lead and the greatest error (12.76%) for the V6 lead. The RNTE signals show the greatest mean error (55.51%), indicating that more than half of the signals recovered are flawed. The RATE signals show the smallest error (40.82%) for the V2 lead and the greatest error (76.04%) for the V5 lead. For all types of signals, except the RATEC signals, the error rates for the V1 and V2 leads are smaller than those for the V5 and V6 leads.

The Corr value of the RATE signals is "1" for most of the leads and 0.99 for only the V6 lead. However, regardless of the pretreatment method, the Corr value is high (> 0.7). In particular, for all types of signals, except the RNTE signals, the Corr value is > 0.95. The Corr values for the V5 and V6 leads are smaller than those for the V1 and V2 leads.

The mean variation of the Amp value (representing a ratio of the amplitudes) is 0.02 for the RATE signals. The mean variation of the Amp value is highest in the RNTE signals (0.21). In particular, the signal amplitude decreases to 70% for the V1 lead, but increases to 1.5 for the V6 lead in the RNTE signals. For all types of signals, the Amp value for the V6 lead is greater than "1", indicating that the signals recovered for the

TABLE III
RESULT OF QUALITATIVE EVALUATION METRICS

		V1	V2	V3	V4	V5	V6	Mean
PSD [%]	RATE	**4.32**	**5.47**	**4.46**	**8.40**	**7.87**	**12.76**	**7.21**
	RNTE	42.58	40.82	54.73	53.84	76.04	65.07	55.51
	RARE	14.75	27.95	39.34	46.72	33.94	30.44	32.19
	RAOE	6.50	14.47	15.46	20.88	32.74	30.71	20.13
	RATEC	27.65	40.50	40.93	44.82	34.16	33.22	36.88
Corr	RATE	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**0.99**	**1.00**
	RNTE	0.93	0.93	0.91	0.88	0.78	0.70	0.86
	RARE	0.99	0.98	0.96	0.96	0.96	0.96	0.97
	RAOE	**1.00**	0.99	0.99	0.97	0.97	0.95	0.98
	RATEC	0.97	0.96	0.96	0.94	0.95	0.96	0.96
Amp	RATE	0.97	1.01	**0.99**	**0.96**	**0.98**	**1.02**	**0.02**
	RNTE	0.67	0.89	**0.99**	0.86	0.87	1.53	0.21
	RARE	0.94	1.05	**0.99**	0.92	0.95	1.03	0.05
	RAOE	**0.98**	0.93	**0.99**	**0.96**	1.13	1.03	0.05
	RATEC	0.90	**1.00**	0.84	0.61	0.92	1.10	0.14

TABLE IV
STATISTICAL EVALUATION OF QUALITATIVE EVALUATION METRICS USING ONE WAY ANOVA

	PSD				Amp			
	df	MS	F	p-value	df	MS	F	p-value
	RUTE	RARE	RAOE	RATEC	RUTE	RARE	RAOE	RATEC
V1	4	225505.140	5031.157	<0.0001	4	14.590	1503.147	<0.0001
	p<0.05	p<0.05	p<0.05	p<0.05	p<0.05	p<0.05	**p>0.05**	p<0.05
V2	4	218350.812	246.670	<0.0001	4	3.933	54.750	<0.0001
	p<0.05	p<0.05	p<0.05	p<0.05	p<0.05	**p>0.05**	p<0.05	**p>0.05**
V3	4	370593.136	244.215	<0.0001	4	4.085	52.576	<0.0001
	p<0.05	p<0.05	p<0.05	p<0.05	**p>0.05**	**p>0.05**	**p>0.05**	p<0.05
V4	4	330515.993	549.052	<0.0001	4	19.158	290.093	<0.0001
	p<0.05	p<0.05	p<0.05	p<0.05	p<0.05	**p>0.05**	**p>0.05**	p<0.05
V5	4	532263.109	1430.125	<0.0001	4	8.388	127.235	<0.0001
	p<0.05	p<0.05	p<0.05	p<0.05	p<0.05	**p>0.05**	p<0.05	p<0.05
V6	4	3436885.125	5228.081	<0.0001	4	41.644	1066.689	<0.0001
	p<0.05	p<0.05	p<0.05	p<0.05	p<0.05	**p>0.05**	**p>0.05**	p<0.05

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2019.2936583, IEEE Journal of Biomedical and Health Informatics

V6 lead are all amplified. In addition, the PSD may have been greatest for the V6 lead because of this amplification. Similar to the result for the visual inspection, the QEM results also show that the RATE method has the best performance.

Table 4 shows the statistical results for PSD and Amp values. Corr did not perform statistical evaluation because it had high correlation regardless of the preprocessing method. One-way ANOVA tests were performed to verify the significance of each RATE, RNTE, RARE, RAOE, and RATEC scheme for each V lead generated. The PSD and Amp values calculated from the samples generated by each method were used as inputs to the one-way ANOVA test. Since the number of n in each group is the same, Bonferroni post hoc test confirms that RATE is significant compared with other methods (RNTE, RARE, RAOE, RATEC).

Both PSD and Amp were confirmed to be significant by one-way ANOVA test ($p < 0.0001$). The PSD index was significantly different from all other results in the post hoc test

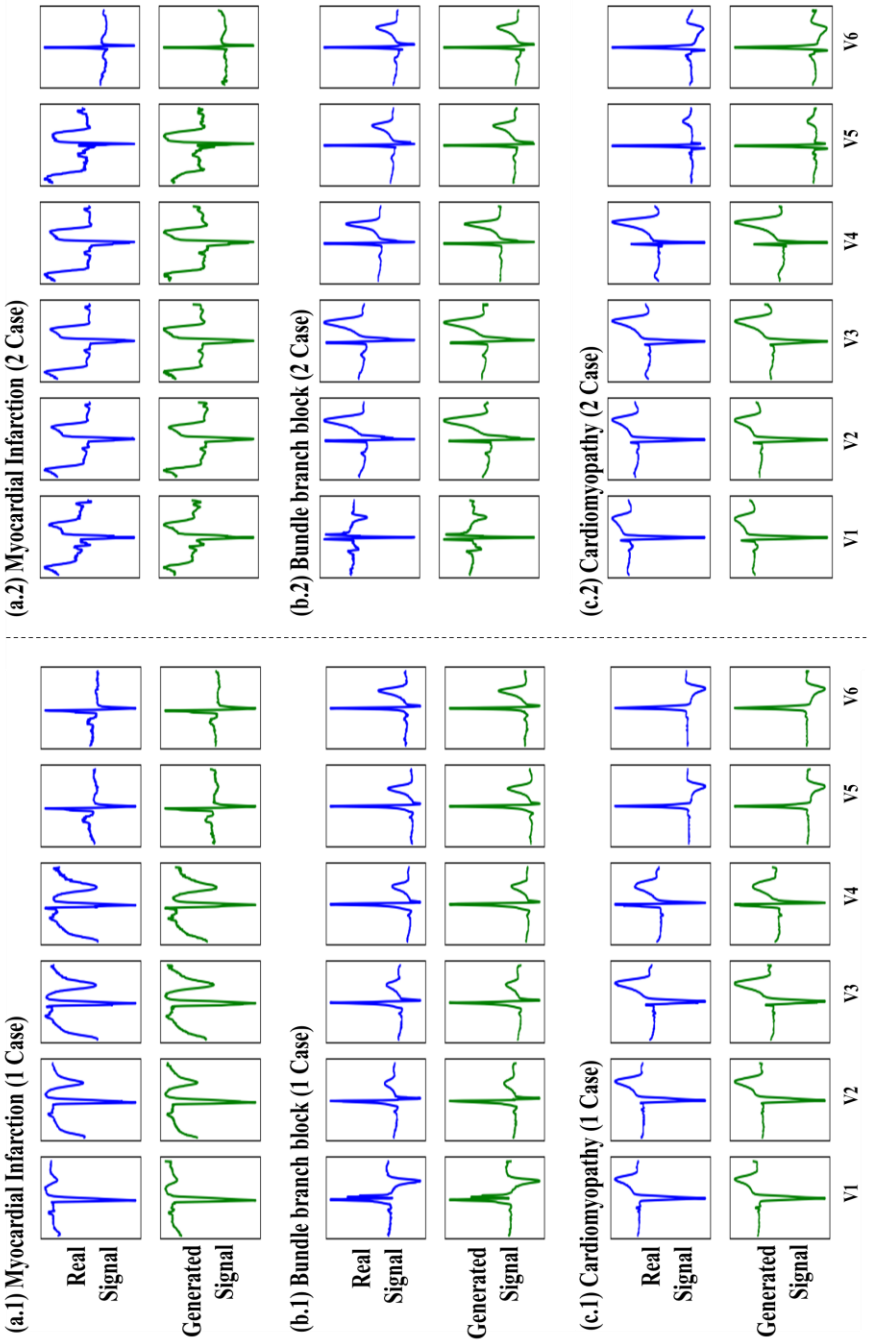

Fig. 4. Two Anomaly Cases of Generated ECG: (a) Myocardial Infraction, (b) Bundle Branch Block, (c) Cardiomyopathy

($p < 0.05$). However, post hoc test results in the Amp index indicate that the RATE does not have a significant difference depending on the leads. Especially, in case of RARE, it was not significant in the case of lead (V2, V3, V4, V5, V6). It shows that Amp can be synthesized similar even if preprocessing is wrong.

E. Case Study

1) Anomaly Cases

For clinical validation, anomaly cases were analyzed using myocardial infraction, bundle branch block, and cardiomyopathy dataset according to the diagnostic class of PTB database.

Figure 4 is a V leads ECG result generated from two data cases for each diagnostic class. (a), (b) and (c) of Figure 4 show the cases of myocardial infraction case, bundle branch block and cardiomyopathy, respectively. Blue is actually measured V lead signals and green color shows generated V lead signals. As a result of visual inspection of the synthesized signals through RATE, even in anomaly cases according to each diagnostic class, the actual signal and the generated signal are generated without a big difference. In particular, when the peak occurs twice as in (a.2), when the peak is broken as shown in (b.1), even if T-wave is inverted such as V5 and V6 in (c.1) and V6 in (c.2), synthesis of V leads performs well.

2) MIT-BIH Arrhythmia Dataset

We used MIT-BIH arrhythmia dataset, another database provided by physionet, for the case study to confirm that the proposed algorithm works well [13]. The MIT-BIH arrhythmia database provides information of two leads. Therefore, clinical validation was performed through the MLII lead and V1 lead provided by the database. The sampling rate of the MIT-BIH dataset is 360 Hz, which is different from the PTB dataset, but down-sampled at 250Hz to apply same preprocessing method. The V1 lead was synthesized after the signals of MLII lead and V1 lead were converted to bi-dimensional space through RATE preprocessing.

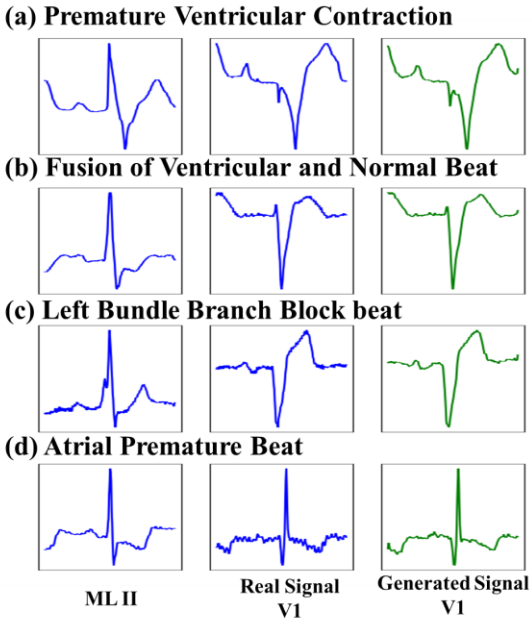

Fig. 5. Generated ECG from MIT-BIH Dataset: (a) Premature Ventricular Contraction, (b)Fusion of Ventricular and Normal Beat, (c) Left Bundle Branch Block Beat, (d) Atrial Premature Beat

Figure 5 is the result of V1 lead ECG generated from four data cases. The (a) is a case of premature ventricular contraction, the (b) is a case of fusion of ventricular and normal beat, the (c) is a case of left bundle branch block beat, and the (d) is a case of atrial premature beat. Blue is the actual measured signal and green indicates the generated signal. As a result of visual inspection of the generated signal, it was confirmed that there is no significant difference between actual and generated signals in various anomaly cases of MIT-BIH arrhythmia dataset. Also, (d) of Figure 5 shows that the noisy component of the original signal is synthesized and reduced.

IV. Discussion

In the present study, V-leads ECG signals were generated from the MLII lead. Recently, various portable ECG hardware devices that allow simple measurement, such as AliveCor and smart watches, have been developed and approved by the US Food and Drug Administration (US-FDA). The portable ECG devices are used for measurement at the limbs, usually for heart-rate monitoring at home [21]. The limb ECG includes the MLI, MLII, and MLIII methods to measure different body parts. The MLII method, which shows the P-wave most significantly, gives the typical ECG waveform [1]. In addition, the MLII signals are the representative ECG signals that are used for the diagnosis of arrhythmia and are included in the open arrhythmia data provided by PhysioNet [13]. However, the MLII signals are difficult to apply to other areas except for home healthcare monitoring, because they do not provide sufficient ECG information. In addition, because the heart has a three-dimensional structure, the V-leads signals are necessary to accurately understand cardiac motion [3]. However, the V-leads ECG method requires that many leads should be attached to the chest, so the measurement is inconvenient. Therefore, the recovery of the V-leads ECG signals from the MLII lead is important for overcoming the limitations of limb-lead ECG applied only for heart-rate monitoring.

In the R-peak-aligned GAN, the excellent performance of GAN for image translation is applied to the time domain. Because the ECG signals have a constant pattern of P-Q-R-S-T waves, it is important to reconstruct the patterns of the ECG signals accurately. The R-peak alignment is the process by which the ECG signals are synchronized with reference to the R-wave, to retain the semantic information contained in the ECG signals. The average heart rate is 70 / min. The physiological components of ECG signals may be recovered by segmenting one second of the signals with reference to the R-peak. However, it is difficult to recover the continuous P-Q-R-S-T pattern when using middle point between peaks. In particular, if the heart rate is less than 60 / min, it may be difficult to extract other ECG parameters as well as the QRS duration and QT interval. If the heart rate is less than 60 / min, R-peak alignment before and after 500 ms with reference to the R-wave can recover the typical ECG pattern of the ECG within the segmented epoch. The importance of R-peak alignment may be confirmed by the RNTE results obtained without the R-peak alignment. From visual inspection, the pattern trend of the ECG signals that have undergone the R-peak alignment may be

recovered, but those that have not undergone R-peak alignment appear randomized. In addition, the ECG signals that have not undergone R-peak alignment have lost their semantic information, and show the lowest performance in terms of all the indices except the IS. The IS of the ECG signals that have not undergone R-peak alignment may be misrepresented in the inception net, because a single uniform image generated, lacks the particular patterns of the ECG signals.

Ordered time-sequence embedding plays the role of converting the time domain signals while retaining the time-dependent information. The ECG signals have a time-dependent flow and reflect sequential information related to previous events. Therefore, alignment of the data according to the time sequence is very important. In the present study, the RARE and RAOE methods were employed to verify the time sequence. In both of these methods, the ECG pattern was retained through R-peak alignment. The RARE signals include the flow of time, while the RAOE signals do not reflect information about the flow of time. With visual inspection, the P-Q-R-S-T patterns of the ECG signals are recovered in both methods, but the small signal flows may not be recovered. This may even result in the swapping of amplification between the Q-wave and the S-wave. In addition, as in the case of the V5 lead (shown in Figure 3e.1); a large error may arise from time-to-time between the original signals and the recovered signals. Because the ordered time sequence transmits the phasic information depending on the data flow, the GAN may learn the long-term information of the ECG signals. Through this, the flows of tiny signals may be recovered using the ordered time sequence.

Finally, the learning by the GAN was performed by pairing each of the V leads with an MLII lead. The paired dataset allows generation through the exact mapping of the lead-ECG signals that reflect individual differences. It is important to retain the individual properties of the ECG signals, because the ECG signals obtained from different individuals have different features. To verify this, the cycleGAN was employed in the present study. The cycleGAN, which has the GAN learn through clusters, does not allow 1-to-1 mapping [17]. With visual inspection, the RATEC method may give seemingly acceptable V-leads ECG signals. However, in comparison with the real data, the error of the RATEC signals is greater than that of the ECG signals obtained using the pair dataset. The IS and SSIM values of the generated signals (shown in Table 1) are small because the RATEC signals were intensively generalized and thus had low diversity. Table 3 shows that the RATEC signals have a high Corr value, because of the high generalization performance of the method, but their PSD and Amp values, directly related to the real signals, are low. Therefore, when the cycleGAN is employed, the generalization performance may be good, but the individual differences may not be well reflected. These results showed that the limb-lead signals contain the components that reflect the individual characteristics. Therefore, in leads data reconstruction, the individual properties may be retained by using the dataset of paired leads.

In the present study, V-leads ECG data synthesis was performed through R-peak alignment, ordered time-sequence embedding, and use of a paired dataset. Table 5 shows the overview of previous studies related to ECG lead synthesis [2], [4], [5], [22], [23], [24], [25], [26]. Previous studies have been performed to synthesize standard 12 leads from arbitrary 3 leads or to synthesize Frank vectorcardiography (VCG). Also, vector coefficient calculation based on ECG theory and PCA were used for the ECG reconstruction in the past. The accuracy

TABLE V
AN OVERVIEW OF THE STUDIES RELATED TO ECG LEAD SYNTHESIS

Author	Purpose	Database	Method	Result
S. Maheshwari et al [2].	3 Lead -> 12 Lead	PTBDB	Heart–Vector Projection Theory	R2: 91.87%
D. Dawson et al [4].	12 Lead -> Frank VCG Frank VCG -> 12 Lead	PTBDB	Affine Transform Dower Transform	R2(Affine): 92.7% (12->3), 84.79% (3->12) R2(Dower): 75.60% (12->3), 77.31% (3->12)
S. Maheshwari et al [5].	12 Lead -> Frank VCG	PTBDB CSEDB	Principal Component Analysis	R2: 73.7% Corr: 0.869
D.M. Schreck el al [19].	3 Lead -> 12 Lead	PTBDB MRMC	Nonlinear Optimization Model	Corr: 0.867
Guldenring et al [20].	12 Lead -> Frank VCG	BSPMs	Multivariate Linear Regression	RMSE: 28.3(x), 45.6(y), 37.4(z)
Vullings et al [21].	12 Lead -> VCG	MIT/BIH PTBDB	Bayesian Vectorcardiography	MSE: 17%
HADZIEVSKI et al [22].	3 Lead -> 12 Lead	Own	Patient Transformation Matrix	ST segment deviation >= 0.1 mV
Atoui et al [23].	3 Lead -> V Lead	CSEDB	Neural Network	Corr: 0.932

of the results obtained in the present study using the R-peak aligned GAN to synthesis V leads from MLII lead were about 93%, and those of previous studies conducted using the PTB dataset were 73% to 94%. In previous studies where the accuracy was about 94%, the lead synthesis was performed by using the V2 lead to calculate the coefficients of the other V leads. However, in these previous studies, with increase of the n value, the accuracy decreased, and the calculation complexity increased because individual calculations were required. The GAN allows for non-linear transformation and learning with many samples; thus, generalization may be performed. Therefore, the R-peak-aligned GAN that has undergone learning in the present study is reusable.

This study confirmed feasibility by synthesizing V leads signal from limb lead which can be measured simply for portable ECG device application. In addition, a case study was conducted using the PTB database and the clinical dataset of the MIT-BIH database provided by physionet. Clinical validation of cardiac disease cases confirmed that lead synthesis was similar to real signal for various cardiac diseases and different databases such as myocardial infraction, bundle branch block, and cardiomyopathy. When the heart rate is less than 60 / min, the P-Q-R-S-T pattern can detect ECG abnormalities. However, it is difficult to know the detailed cause of cardiac abnormality in the long heart rate interval such as heart rate which is more than 60 / min. In particular, the first degree heart block has a PR interval of more than 200 ms or a QT interval of up to 590 ms. So, it could be difficult to recover the overall P-Q-R-S-T pattern according to interval duration [27]. Also, it is difficult to know exact clinical meaning only using V lead synthesis. Therefore, it is necessary to further investigate whether ECG signal recovery is successful whether other leads can be translated well to generate a standard 12 - lead ECG as in the previous studies in Table 5. In addition, this study has limitations in that it is designed for open datasets. The open dataset is acquired in a stable environment in the hospital, so the signal length is short, and the effect of noise is low. However, when measuring ECG in real time, the heart rate changes dynamically depending on the activity. For practical portable device application, it is necessary to apply leads synthesis from data obtained from actual portable ECG besides open dataset. Also, portable ECG is highly affected by noise to the surrounding environment. Therefore, it is important to study de-noising technique to apply it to actual environment. In previous studies, there have been various studies on ECG noise reduction using GANs, RNNs, etc. [6], [28], [29]. In particular, when long short-term memory models (LSTM) is applied, root mean square error is about 0.25-0.35, which is excellent [29]. Therefore, further studies on ECG de-noising based on these existing studies are needed, and it is thought that the performance of R-peak aligned GAN can be further enhanced through the further study.

V. Conclusions

In the present study, V-leads ECG signals were generated through the GAN from the ECG signals obtained from the MLII lead, which is a limb lead. The present study is different from previous studies because R-peak alignment, ordered time-sequence embedding, and a paired dataset were employed. Current portable ECG devices may not be applied clinically, because only limited ECG signals obtained from the limb lead are provided. In addition, such methods as R-peak detection and S-T segment analysis are used to develop ECG applications. Therefore, the ECG pattern reconstruction accuracy is critical to the application of ECG signals. The R-peak-aligned GAN investigated in the present study may allow for the recovery of the lead-ECG signals through the portable ECG lead so that the recovered data may be used both as a mobile environment and for hospital clinical data. Furthermore, the one-to-multi lead reconstruction may help to develop advanced portable ECG hardware and overcome the problems involved with the inconvenience of attaching many leads and the insufficiency of data storage space.

Acknowledgement

This work was supported by the Bio & Medical Technology Development Program of the National Research Foundation(NRF) funded by the Ministry of Science and ICT.(NRF-2018M3A9H6081483)

References

[1] J. Francis, “ECG monitoring leads and special leads,” *Indian Pacing Electrophysiol. J.*, vol. 16, no. 3, pp. 92-95, 2016.

[2] S. Maheshwari, et al., “Accurate and reliable 3-lead to 12-lead ECG reconstruction methodology for remote health monitoring applications,” *IRBM* vol. 35, no. 6, pp. 341-350, 2014.

[3] M. Vozda, and M. Cerny, “Methods for derivation of orthogonal leads from 12-lead electrocardiogram: A review,” *Biomed. Signal Process. Control*, vol. 19, pp. 23-34, 2015.

[4] D. Dawson, H. Yang, M. Malshe, S. T. Bukkapatnam, B. Benjamin, and R. Komanduri, “Linear affine transformations between 3-lead (Frank XYZ leads) vectorcardiogram and 12-lead electrocardiogram signals,” *J. Electrocardiol.*, vol. 42, no. 6, pp. 622-630, 2009.

[5] S. Maheshwari, A. Acharyya, M. Schiariti, and P. E. Puddu, “Frank vectorcardiographic system from standard 12 lead ECG: An effort to enhance cardiovascular diagnosis,” *J. Electrocardiol.*, vol. 49, no. 2, pp. 231-242, 2016.

[6] P. Xiong, et al., “ECG signal enhancement based on improved denoising auto-encoder,” *Eng. Appl. Artif. Intell.*, vol. 52 pp. 194-202, 2016.

[7] D. Rajan, and Jayaraman J. Thiagarajan, “A Generative Modeling Approach to Limited Channel ECG Classification,” *arXiv*, vol. 1802.06458, 2018.

[8] T. Golany, & Radinsky, K. , “PGANs: Personalized Generative Adversarial Networks for ECG Synthesis to Improve Patient-Specific Deep ECG Classification.,” *Association for the Advancement of Artificial Intelligence*, 2019.

[9] E. Brophy, Wang, Z., & Ward, T. E., “Quick and Easy Time Series Generation with Established Image-based GANs,” *arXiv preprint*, 2019.

[10] J. Johnson, Alexandre Alahi, and Li Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," *European Conference on Computer Vision*, vol. Springer, 2016.

[11] I. Goodfellow, et al, "Generative adversarial nets," *Adv. Neural Inf. Process. Syst.*, 2014.

[12] P. Isola, et al, "Image-to-image translation with conditional adversarial networks," *arXiv*, 2017.

[13] A. L. Goldberger, L. A. 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, pp. E215-220, 2000.

[14] I. Amer-Wahlin *et al.*, "Fetal electrocardiography ST analysis for intrapartum monitoring: a critical appraisal of conflicting evidence and a way forward," *Am. J. Obstet. Gynecol.*, 2019.

[15] J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE Trans. Biomed. Eng.*, vol. 32, no. 3, pp. 230-236, 1985.

[16] K. He, et al, "Deep Residual Learning for Image Recognition," *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016.

[17] J.-Y. Zhu, et al, "Unpaired image-to-image translation using cycle-consistent adversarial networks," *arXiv* 2017.

[18] A. Borji, "Pros and Cons of GAN Evaluation Measures," *arXiv*, vol. 1802, no. 03446 2018.

[19] K. R. Rao, and H. R. Wu, "Structural similarity based image quality assessment," *Digital Video image quality and perceptual coding, CRC Press*, pp. 261-278, 2005.

[20] B. Singh, Amandeep Kaur, and Jugraj Singh, "A Review of ECG Data Compression Techniques," *Int. J. Comput. Appl.*, vol. 116, no. 11, 2015.

[21] D. E. Albert, et al, "Methods and systems for cardiac monitoring with mobile devices and accessories," *U.S. Patent Application*, vol. 14, no. 692, pp. 563.

[22] D. M. Schreck and R. D. Fishberg, "Derivation of the 12-lead electrocardiogram and 3-lead vectorcardiogram," *Am. J. Emerg. Med.*, vol. 31, no. 8, pp. 1183-1190, 2013.

[23] D. Guldenring, D. D. Finlay, D. G. Strauss, L. Galeotti, C. D. Nugent, M. P. Donnelly, and R. R. Bond, "Transformation of the Mason-Likar 12-lead electrocardiogram to the Frank vectorcardiogram," *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, vol. 2012, pp. 677-680, 2012.

[24] R. Vullings, C. H. Peters, I. Mossavat, S. G. Oei, and J. W. Bergmans, "Bayesian approach to patient-tailored vectorcardiography," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 3, pp. 586-595, 2010.

[25] L. Hadzievski, B. Bojovic, V. Vukcevic, P. Belicev, S. Pavlovic, Z. Vasiljevic-Pokrajcic, and M. Ostojic, "A novel mobile transtelephonic system with synthesized 12-lead ECG," *IEEE Trans. Inf. Technol. Biomed.*, vol. 8, no. 4, pp. 428-438, 2004.

[26] H. Atoui, J. Fayn, and P. Rubel, "A novel neural-network model for deriving standard 12-lead ECGs from serial three-lead ECGs: application to self-care," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 3, pp. 883-890, 2010.

[27] T. Kasar, H. C. Kafali, and Y. Ergul, "A rare association: first degree AV block and long QT syndrome," *Cardiol. Young*, pp. 1-2, 2019.

[28] J. Wang, et al, "Adversarial de-noising of electrocardiogram," *Neurocomputing*, vol. 349, pp. 212-224, 2019.

[29] J. Guan, et al, "Automated Dynamic Electrocardiogram Noise Reduction Using Multilayer LSTM Network," *Proceedings of the 15th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services. ACM*, 2018.