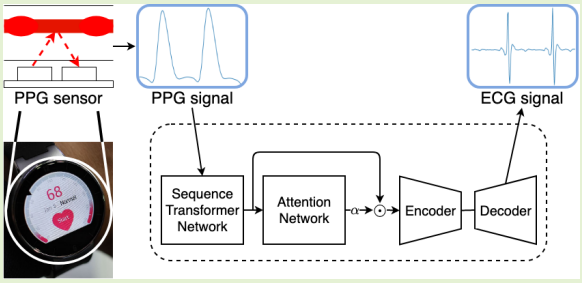
Reconstructing QRS Complex from PPG by Transformed Attentional Neural Networks

Hong-Yu Chiu, Hong-Han Shuai, *Member, IEEE*, Paul C.-P. Chao, *Senior Member, IEEE*

***Abstract*— Technology that translates photoplethysmogram (PPG) into the QRS complex of electrocardiogram (ECG) would be transformative for people who require continuously monitoring. However, directly decoding the QRS complex of ECG from PPG is challenging because PPG signals usually have different offsets due to 1) different devices, and 2) personal differences, which makes the alignment difficult. In this paper, we make the first attempt to reconstruct the QRS complex of ECG only from the recording of PPG by an end-to-end deep learning-based approach. Specifically, we propose a novel encoder-decoder architecture containing three components: 1) a sequence transformer network which automatically calibrates the offset, 2) an attention network, which dynamically identifies regions of interest, and 3) a new QRS complex-enhanced loss for better reconstruction. The experiment results on a real dataset demonstrate the effectiveness of the proposed method: 3.67% R peak failure rate of the reconstructed ECG and high correlation of pulse transit time between the reconstructed QRS complex and the groundtruth QRS complex ($\rho = 0.844$), which creates a new opportunity for low-cost clinical studies via the waveform-level reconstruction of the QRS complex of ECG from PPG.**



***Index Terms*— Convolutional neural network, electrocardiography, encoder-decoder, photoplethysmography, transform network.**

I. Introduction

WITH the increase in world population along with a significantly aging subset, the continuous monitoring of vital signs from electrocardiogram (ECG) has become more and more important for personal healthcare. Specifically, ECG contains five peaks (represented by the letters P, Q, R, S and T) which reflects the electrical activity of the heart by exploiting electrodes placed on the skin, and thus providing essential information for cardiovascular pathology [1]. However, it is inconvenient to measure ECG by the standard 12-lead ECG device since it requires putting several electrodes on different skin positions, and this may cause skin irritation and discomfort during the recording.

On the other hand, photoplethysmogram (PPG) is an optically obtained signal that can be used to detect blood volume changes in the microvascular bed of tissue [2]. Moreover, the duration, magnitude and shape characteristics of the PPG waveform can be translated to blood oxygen saturation [3], heart rate [4], blood pressure [5], cardiac output [6] and respiration rate [7]. Compared to ECG, the process of deriving PPG is noninvasive, more convenient to set up, and low-cost. For example, some consumer-grade wearable devices like smartwatches and smartphones have photo-transistors that can offer continuous and long-term monitoring.

Although PPG technology has become popular for healthcare monitoring [8], ECG is still the standard and essential measurement for medical diagnosis with abundant supporting literature and studies. As such, doctors still rely on ECG for diagnosis instead of using PPG. However, PPG and ECG are intrinsically correlated since the changes in blood volume are influenced by the electrical activity of the heart. The peak to peak interval of PPG is also known to be highly correlated with the R-R interval (the time elapsed between two successive R peaks) of an ECG, indicating the possibility of deriving ECG from PPG. Therefore, based on these observations, we propose to utilize this correlation for directly reconstructing ECG from the PPG measurement, which can leverage the advantages of both the low cost and easy accessibility of PPG in addition to the well studied base of ECG.

In light of the recent breakthroughs in the deep learning field that opens up new possibilities for signal processing, we propose in this study an end-to-end deep learning-based approach to reconstruct the ECG waveform using the PPG measurement. One of the possible solutions is to directly use a Recurrent Neural Network (RNN) to transform PPG to ECG. However, the RNN is sensitive to signal noises and computed with time dependency, which may not be suitable for real-time applications on resource-limited devices, such as ASIC

H.-Y. Chiu is with the BASIC Lab, Department of Electrical and Computer Engineering, National Chiao Tung University (NCTU), Hsinchu, Taiwan, R.O.C. (e-mail: james77777778.eed07g@nctu.edu.tw).

H.-H. Shuai is with the BASIC Lab, Department of Electrical and Computer Engineering, National Chiao Tung University (NCTU), Hsinchu, Taiwan, R.O.C. (e-mail: hhshuai@nctu.edu.tw).

Paul C.-P. Chao is with the Sensors IC Lab, Department of Electrical and Computer Engineering, National Chiao Tung University (NCTU), Hsinchu, Taiwan, R.O.C. (e-mail: pchao@mail.nctu.edu.tw).

and smartwatches. Therefore, we adopted a Convolutional Neural Network (CNN) based encoder-decoder architecture as a transform network to reconstruct ECG for two reasons: 1) the CNN is highly parallel, and 2) the encoder-decoder architecture has been proved to be successful for tasks of domain adaptation [9], signal denoising [10]–[12] and anomaly detection [10]. Further, to deal with the inputs from different instances, a sequence transformer network is proposed for the temporal and magnitude transformation of the inputs. We also used an attention network to make the transform network focus on the correct part of PPG to better reconstruct the ECG waveform. The contributions of this study are in the following.

To the best of our knowledge, the proposed method is the first end-to-end learning method to reconstruct ECG directly from the raw PPG recordings without any extra handcrafted clinical features or explicit assumptions, which can facilitate various applications of sensors, e.g., continuous monitoring of ECG, Pulse Transit Time (PTT) estimation by using PPG only (can be directly used to measure the blood pressure).

We propose a novel encoder-decoder architecture containing two components: 1) a sequence transformer network which automatically calibrates the offset, and 2) an attention network, which dynamically identifies regions of interest. Moreover, a new loss function for better alignment of R peaks is also proposed, while a signal attention is further added to make the transform network focus on the correct part of PPG to better reconstruct the ECG waveform.

The experiment results on the real dataset show that the success ratio of R peak detection from the reconstructed ECG waveform is 96.33%, and the difference between the reconstructed R peak index value and the groundtruth index value is only 16.11ms on average. In addition, the correlation between the PPT calculated by the input PPG and the reconstructed ECG and the PTT calculated by the input PPG and the groundtruth ECG is high (= 0.844), which indicates the potential for clinical application.

The remainder of this paper is structured as follows. Section II presents the related works including a review of the literature as well as current commercial products. In Section III, we describe the database and propose a novel transformed attentional model for ECG reconstruction. We introduce our test scenarios and present our experimental results in Section IV. Finally, Section V concludes this paper.

II. Related Works

Several commercial products aim to provide ECG recordings, such as ECG watches like the Apple Watch Series 4/5[1] which require users to put their fingers on the watch to form a closed circuit[2] and cannot, therefore, be used for continuous monitoring. ECG watches also require the additional ECG electrodes, which are more expensive than the proposed solution in this paper. On the other hand, QARDIOCORE[3] and MAX-ECG-MONITOR[4] are belt-like ECG monitors with wearable electrodes that provide continuous monitoring. These devices directly record the ECG signal as long as they are correctly fitted, but the user has to put on a specific device and they also cost a lot compared to PPG monitoring devices. Furthermore, the electrodes of the products might also cause skin irritation and discomfort during the recording. Another wearable ECG product, KardiaMobile[5], provides ECG monitoring by electrodes on small pads which can be attached to phones. However, KardiaMobile is unable to provide long-term ECG monitoring due to requirement of putting the users' fingers on the device to form a closed circuit for capturing the electrical signal.

A recent line of research also aims to provide ECG by wearable devices. For example, in [13], the authors present a fully disposable single-lead ECG patch composed of Ag-AgCl electrodes and components integrated on a flexible circuit board. However, no evaluation of the quality of the ECG recording was presented in the study and the ECG patch only has a 7-day lifespan, which is not suitable for long-term monitoring. [14] introduces a low-power wearable ECG monitoring system which uses a belt-like device with electrodes to form the closed circuit. However, this device also suffers from not only the issues of inconvenience and discomfort for users similar to commercial belt-like products, but also ECG baseline drift due to respiration and movements. In [15], a non-contact wearable wireless ECG system is presented which can measure ECG signals over a textile-based interface material between the skin and electrodes. The electrodes do not have direct contact with the skin, thus preventing any skin irritations or possible allergies. Yet, the authors indicate that the electrodes need to be firmly integrated over the cloth to reduce the displacement, and the electrostatic coupling of the skin electrodes needs to be improved. Further, it is unfortunate that there was no evaluation of the quality of ECG in [15].

In summary, Table I compares the five current technologies for ECG monitoring: 1) finger-based ECG device, 2) ECG watch, 3) belt-like ECG device, 4) standard ECG device, and 5) the system we propose in this paper. A finger-based ECG device (*e.g.*, KardiaMobile) measures lead-I ECG using a small pad with two electrodes, which is similar to an ECG watch (*e.g.*, Apple Watch Series 4/5). These two types of ECG devices are affordable and are easy to use, but neither of them can be used for long-term ECG monitoring since they require users to put their fingers on the devices to form closed circuits. On the other hand, belt-like ECG devices (*e.g.*, QARDIOCORE and MAX-ECG-MONITOR) provide long-term ECG monitoring, but they generate non-standard ECG types and require users to put on a special device that is also more expensive. Standard ECG devices employ the gold standard 12-lead ECG, but they are very expensive and require a complicated installation, which makes them unsuitable for daily monitoring. In contrast, our proposed system does not require any closed circuit to detect electrical

[1] https://www.apple.com/apple-watch-series-5/
[2] https://youtu.be/lpXfQDK_uuw
[3] https://store.getqardio.com/products/qardiocore
[4] https://www.maximintegrated.com/en/products/sensors/MAX-ECG-MONITOR.html
[5] https://www.alivecor.com/kardiamobile

TABLE I
COMPARISON OF DIFFERENT APPROACHES FOR VITAL SIGNS.

	ECG type	Cost	Accessibility	Long-term Monitoring
Finger-based ECG Device	Lead I	Low	High	
ECG Watch	Lead I	Medium	High	
Belt-like ECG Device	Unclassified	Medium	Medium	X
Standard ECG Device	Standard 12-lead	High	Poor	
Ours	Lead II	Low	High	X

activity of the heart, which can be useful for long-term ECG monitoring research since abnormal signals, such as those that may occur in sleep, are rare and difficult to find if we can only measure the ECG waves several times a day. Moreover, our proposed method differs from current wearable single-lead ECG systems insofar as our model learns to reconstruct the standard lead II ECG, while current wearable single-lead ECG devices have to deal with a certain amount of disturbances and uninterpretable waveform, especially during heart rate variability, and also when counting premature atrial contraction and premature ventricular contraction, as has been pointed out in the literature [16], [17].

There are few previous works aiming to reconstruct ECG from PPG. In [18], the author exploits a machine-learning-based feature selection approach to estimate the range of ECG parameters using PPG features. Even though their system achieves more than 90% accuracy in estimating ECG parameters on a benchmark hospital dataset, the lack of complete ECG waveform reconstruction is an obstacle to the wide-scale adoption of their system. [19] proposes method that can reconstruct the whole ECG waveform by first aligning each PPG cycle to ECG cycle and then mapping the discrete cosine transform (DCT) coefficients of each PPG cycle to those of the corresponding ECG cycle. However the cycle-wise alignment and segmentation in the preprocessing stage lose temporal information, such as pulse transit time and heart rate variation, which are important clinical factors. Furthermore, the preprocessing overhead restricts the capacity of their applications and makes it difficult to build a real-time transformation.

III. PROPOSED ALGORITHM

A. Database Description

The public data in this study were provided by the University of Queensland (UQVSD dataset) [20] and come from 32 surgical patients who underwent anaesthesia at the Royal Adelaide Hospital 55 hours of readings, or 4.5 GB data, which are mainly of normal sinus rhythm). PPG and ECG are paired data, which means that they are simultaneously measured for transform learning. We filtered out the noisy data which may be caused by instrumental and personal errors mostly at the head and tail-end of the recordings and this also kept the recording lengths balanced.

Moreover, to evaluate the generalizability of our method, we also conducted experiments on another dataset, BIDMC PPG and Respiration Dataset (BIDMC) [21], which contains paired PPG and ECG measured from 53 critically-ill patients (8-minute duration for each) during their hospital care at the Beth Israel Deaconess Medical Centre. The waveforms of PPG and ECG in this dataset are quite different from UQVSD set because they are more noisy and abnormal. We mainly used UQVSD dataset for evaluation but used BIDMC dataset to study the robustness and generalizability of our algorithm in Section IV-C.

B. Proposed Method

In this section, we present our method to reconstruct the ECG directly from the raw PPG recordings. Let $x_i = (x_{i;1}; \quad ; x_{i;t}; \quad ; x_{i;T_i})$ and $y_i = (y_{i;1}; \quad ; y_{i;t}; \quad ; y_{i;T_i})$ respectively denote the PPG and ECG of user i with the total number of time steps T_i, where $x_{i;t}$ and $y_{i;t}$ respectively denote the PPG and ECG of user i at time t. Given a dataset D of simultaneously recorded PPG and ECG with N users, *i.e.*, $D = f(x_i; y_i)j1 \quad i \quad Ng$, we aim to transform x_i into the reconstructed ECG, denoted as $\hat{y}_i$, such that $\hat{y}_i$ is close to y_i for all i. To reconstruct ECG from PPG, we propose a new framework which adopts an encoder-decoder architecture and is further improved by means of three modules, *i.e.*, the sequence transformer network, the attention network and the proposed QRS complex-enhanced loss. Fig. 1 shows the flowchart of the proposed method. Specifically, we adopted the encoder-decoder architecture which learns a shared representation of PPG and ECG by encoding the input PPG to a latent feature and by decoding from the latent feature to construct the ECG. The encoder-decoder architecture can be regarded as a dimension reduction process, and is thus more robust to noises. Moreover, we employed one-dimensional CNN for the encoder and decoder, instead of the RNN, due to the efficiency and robustness of CNNs.

However, the CNN-based encoder-decoder architecture is sensitive to the magnitude and phase of inputs. For example, different PPG sensors may show different offsets, which results in the difficulty of the ECG reconstruction. Inspired by [22], we propose that possible alignment can be learned directly from the inputs by the sequence transformer network, denoted as $STN()$. Fig. 2 shows the architecture of the sequence transformer network which exploits two learnable parameters and to respectively transform the phase and adjust the magnitude of the input. Specifically, the sequence transformer network learns the temporal transformation using , which transforms the original input $x_{i;t}$ and maps it to a new temporal location in the output, *i.e.*,

$$t' = \begin{bmatrix} t \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} t \\ 1 \end{bmatrix} : \tag{1}$$

In other words, (1) maps the original time point t to the transformed time point t' by $_1$ and $_0$, *i.e.*, $x_{i;t'} = x_{i;(\ _1t+\ _0)}$. Moreover, because $t' = \ _1t+\ _0$ is not guaranteed to a positive

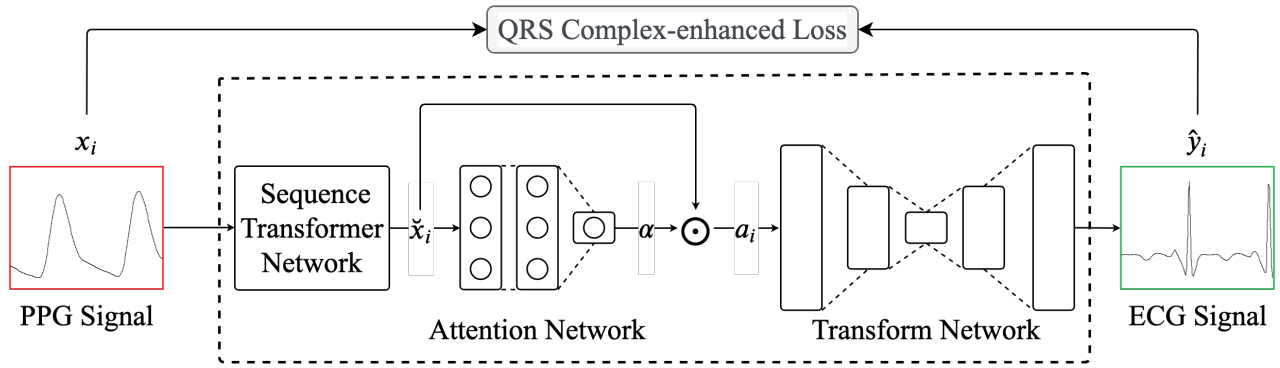


Fig. 1. The architecture of the proposed method.

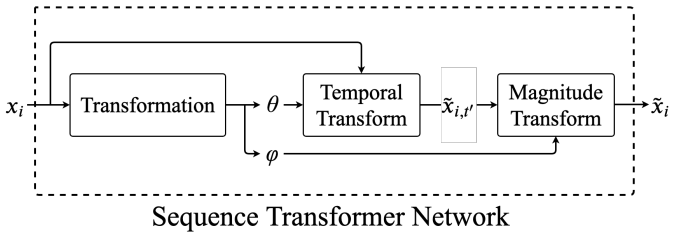


Fig. 2. The architecture of the Sequence Transformer Network.

integer (index), we used a linear sampling to extract the value of the index by taking the average over the two neighbors from the transformed points. For example, if t^0 is 1.5, we calculate $\tilde{x}_{i;t^0}$ by taking the average of $x_{i;1}$ and $x_{i;2}$. After the temporal transformation, we derive the temporally transformed signal $\tilde{x}_i$, and further use the magnitude transformation to adjust the magnitude. Let $x_{i;t}$ denote the magnitude-transformed signal, which can be derived as follows.

$$x_{i;t} = \begin{bmatrix} \tilde{x}_{i;t} \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha_1 & \alpha_0 \end{bmatrix} \begin{bmatrix} \tilde{x}_{i;t} \\ 1 \end{bmatrix}: \tag{2}$$

Equation (2) maps the magnitude of $\tilde{x}_{i;t}$ to $x_{i;t}$ at time t with α_1 and α_0 such that the magnitude can be shifted and scaled, *i.e.*, $x_{i;t} = \alpha_1 \tilde{x}_{i;t} + \alpha_0$. In summary, the temporal and magnitude transformation can be regarded as the normalization for the network, which significantly improves the performance when the inputs are from different users or with the offsets resulting from different sensors. The sequence transformer network can be written as

$$x_i = STN(x_i): \tag{3}$$

Afterward, the attention network, denoted as $Attn()$, is utilized to learn the important part of the input PPG for reconstructing the ECG. We used two fully-connected layers and a softmax function as the activation function, which takes the inputs and normalizes them into a probability distribution.[6] The input x_i is first fed into the attention network to generate an attention weight vector, denoted as α, which has the same dimension as x_i. After that, the attention weight vector is amplified by the input size T because the original attention weight vector is passed through the softmax function which represents the probability of the importance, which may result in gradient vanishing problem. Finally, the amplified attention weight vector α is multiplied by the original inputs in an element-wise manner to generate an attended signal, denoted as a_i, in the following.

$$a_i = T \quad Attn(x_i) \quad x_i; \tag{4}$$

Finally, an encoder and a decoder, denoted as $Encoder()$ and $Decoder()$ respectively, are utilized to transform a_i and together are called the transform network. We used a shallow version of the generative network described in [24], which has been proven to be powerful for speech denoising. Specifically we use the same kernel size of 31, similar to [24], but with a smaller number of filters and only 10 layers instead 22 layers because the sampling rate is smaller, which results in a smaller input size ($T = 200$). In the encoding stage, the input signals are projected and compressed through a number of strided convolutional layers followed by parametric rectified linear units (PReLUs). Our transform network only decreases the resolution twice using strided convolutions to $\frac{1}{4}$ of the original size ($T = 200$ to 50) as it has been proved to get clearer outputs [25]. The decoding stage is the reversed version of the encoding network by means of strided transposed convolutions followed by PReLUs, *i.e.*,

$$\hat{y}_i = Decoder(Encoder(a_i)): \tag{5}$$

It is worth noting that one important feature of the proposed method is its end-to-end training with the ability to deal with raw input signals, which does not require complicated signal processing and handcrafted hyper-parameters to complete the reconstruction. To minimize the distance between the reconstructed ECG and the groundtruth ECG, the L_1 norm is chosen as the loss function, as it is effective for generating sharper signals than the L_2 norm [24].

$$L_1 = \sum_{i2N} jy_i \quad \hat{y}ij: \tag{6}$$

In the following, we introduce the modified L1 loss for a better reconstruction of the QRS complex.

[6]It is worth noting that the attention network used here is the inner-attention that aims at finding the more important part of the inputs [23]. One alternative is to use the self-attention network on the latent vector of the encoder output, but then the number of learning parameters significantly increases, as does the inference time.

C. QRS Complex-Enhanced Loss

Previous work has shown that the QRS complex is more important than P peaks and T peaks, which has been demonstrated to be vital for further pathology studies, *e.g.*, arrhythmic events [26]. However, (6) does not emphasize the importance of the QRS complex. Moreover, the magnitude of R peaks in the reconstructed ECG tends to be smaller than the real value, which leads to the missing detection of R peaks in the ECG analysis. This is because the L1 loss significantly increases when the positions of R peaks are slightly misaligned; the magnitude, therefore, is made smaller to avoid these types of large error. To overcome this issue, we designed a new QRS complex-enhanced loss, L_{QRS}, which encourages the whole neural network to focus on the QRS complex area by an non-normalized Gaussian weighting function using the R peak index as the mean and a hyper-parameter as the variance[7]. Let $c_i = (c_{i;1}; \quad ; c_{i;k}; \quad ; c_{i;K_i})$ denote the R peak index for y_i, where K_i represents the number of R peaks. The QRS complex-enhanced loss is derived as follows.

$$L_{QRS} = \sum_{i=1}^{N} \sum_{t=1}^{T} |y_{i;t} \quad \hat{y}_{i;t}| (1 + \sum_{k=1}^{K_i} e^{\frac{(t \quad c_{i;k})^2}{2^{\ 2}}}); \quad (7)$$

where is a hyper-parameter controlling the impact of the Gaussian weighting function. In summary, if the output $\hat{y}_i$ fails to accurately reconstruct the groundtruth signals y_i in R peak region, the QRS complex-enhanced loss becomes greater. To label R peaks for training, we used the state-of-the-art method, the Hamilton Segmenter [27], which is introduced by a public toolbox, namely BioSPPy [28], for biosignals processing. The Hamilton Segmenter has proven to be a strong classifier to detect the QRS complex with sensitivity of 93% on the MIT/BIT and AHA databases, and it achieves high performance for R peak detection in ECG.

IV. Experiment Results

A. Experiment Setup

Since the goal is to reconstruct ECG from PPG, in the preprocessing stage for both the UQVSD and BIDMC datasets, a min-max scalar normalization is used to transform the magnitude of both signals to $[\ 1; 1]$, while two bandpass filters are applied to ECG and PPG respectively to reduce noise. In accordance with the sampling rate of 100 Hz (UQVSD dataset) and 125 Hz (BIDMC dataset), the size of the sliding window T is respectively set as 200 and 256 to serialize the data. The model is trained for 300 epochs with Adam [29] at a learning rate of 0.0001 and batch size of 256.

Here, we used the data augmentation to effectively learn the transformation starting from any position of PPG. Specifically, we randomly shifted the phases of the input PPG and corresponding groundtruth of ECG to train our model within $[\ 10; 10]$. Therefore, the training task can synthesize the ECG starting from any position. The data augmentation mitigates the boundary issue of the convolutional layer when the target has R peaks at the boundaries, which also decreases the failure probability for the Hamilton Segmenter at the boundaries due to a lack of complete QRS complex. Finally, we used 10-fold cross-validation and calculated the average results.

[7]The setting of will be discussed later in Section IV-E.

B. Evaluation Metrics

To evaluate the model performance, we employed the L1 loss of QRS complex area ($L1_{QRS}$), the L1 loss of non-QRS complex area ($L1_{nQRS}$), the Normalized Mean Absolute Error (NMAE) and the Normalized Root Mean Square Error (NRMSE) as the metrics. The NMAE and NRMSE are calculated as follows.

$$\text{NMAE} = \frac{\sum_{i2N} |y_i \quad \hat{y}_i|}{\sum_{i2N} |y_i|}; \quad (8)$$

$$\text{NRMSE} = \frac{\frac{\sum_{i2N} |y_i \quad \hat{y}_i|^2}{N}}{\max_{i2N} y_i \quad \min_{i2N} y_i}; \quad (9)$$

NMAE and NRMSE are resistant to outliers and can be viewed as overall performance. When identifying the QRS complex areas, previous studies [30] show that the normal intrinsicoid deflection (Q-R interval) is at most 50ms, and the QRS duration is at most 120ms. Therefore, we labeled the range within [-50ms, +70ms] at the R peak index as the QRS complex area and the remaining area as the non-QRS complex area. Moreover, we used three more metrics to show the performance in the R peak area. Let D_{test} denote the test data, and $\hat{c}_{i;k}$ denote the R peak position of $\hat{y}_i$ that is the closest to $c_{i;k}$. The first metric is the Mean Location Error (MLE) of the R peaks, which is measured the difference of the R peak positions between the groundtruth ECG and the reconstructed ECG. The MLE is calculated as follows.

$$\text{MLE} = \frac{1}{|D_{test}|} \frac{1}{K_i} \sum_{(x_i;y_i)2D_{test}} \sum_{k=1}^{K_i} f(c_{i;k}; \hat{c}_{i;k}); \quad (10)$$

$$f(c_{i;k}; \hat{c}_{i;k}) = \begin{cases} |c_{i;k} \quad \hat{c}_{i;k}| & \text{if } |c_{i;k} \quad \hat{c}_{i;k}| < 10 \\ 10 & \text{otherwise.} \end{cases}$$

Second, the Mean Magnitude Error (MME) of the R peaks is proposed to measure the difference of the R peak magnitudes between the groundtruth ECG and the reconstructed ECG. The MME is formulated as follows.

$$\text{MME} = \frac{1}{|D_{test}|} \frac{1}{K_i} \sum_{(x_i;y_i)2D_{test}} \sum_{k=1}^{K_i} |y_{i;c_{i;k}} \quad \hat{y}_{i;c_{i;k}}|: \quad (11)$$

These two metrics reflect the quality of the R peaks, which are vital for detecting beat segments, or any other morphological parameters of the ECG [27]. The MLE of the R peaks indicates how accurate the R peaks are, and the MME of the R peaks represents the magnitude error of the R peaks in the reconstructed ECG. Moreover, the variance of both MLE and MME is also calculated to show the stability of the reconstructed ECG. Finally, we ran the R peak detection algorithm for both the groundtruth and reconstructed ECG and calculated the R peak failure rate (R_{FAIL}) to evaluate the

quality of QRS complex and the adaptability of the systems as follows.

$$R_{FAIL} = \frac{\text{failed R peak detection}}{\text{all R peaks in groundtruth signals}}: \quad (12)$$

It is worth noting that if there is no R peak detected in the reconstructed signals between -75ms and +75ms at the index of the groundtruth R peak, it is considered a failed R peak detection [31].

C. Quantitative Results

To show the improvement of our proposed method, we compared the proposed model with the vanilla LSTM-based transform network (Vanilla LSTM) used in [32] for the speech synthesis[8]. Table II indicates that the proposed model outperforms the Vanilla LSTM for every metric. This is because the CNN better captures the temporal patterns while both the sequence transformer network and attention network help the main transform network to reconstruct the ECG more precisely. In addition, the non-QRS complex area has a much lower L1 loss than the QRS complex area has; this is because the waveform of the QRS complex area is more complicated that of the non-QRS complex area that includes some steady flatlines. The MLE of our model is 1.611 which is equal to 16.11ms error at the R peak index on average at a sampling rate of 100Hz. The minor errors are acceptable for further analysis such as pulse transit time (discussed below). Moreover, the magnitude of ECG generated by the Vanilla LSTM is smaller at R peaks, which results in a high failure rate of R peak detection. For the proposed model, 3890 R peaks in the groundtruth ECG are successfully detected while 142 R peaks are missing on average, thus resulting in a 3.67% R peak failure rate. The proposed model significantly improves the performance to get a lower R peak failure rate because of the better quality of the QRS complex; we credit this improvement to our implemented sequence transformer network, attention network and QRS complex-enhanced loss.

From the efficiency evaluation of the proposed model, Table III presents the average running time of reconstructing the ECG per 1-sec PPG segment with different batch sizes of the input PPG, as run on an Nvidia GTX1070 GPU. The results shows that the proposed model requires 1.514ms to 2.054ms to transform a 1-sec PPG segment. As the batch size increases, our model better utilizes the parallel computing of the GPU since multiple reconstructions can be generated on a batch of inputs. A larger batch size (100) can reduce almost 25% of the inference time compared to only one input at a time. Compared to the Vanilla LSTM, the running time of the proposed model is much faster with a larger batch size since it can calculate the convolution of the same filter on multiple locations in parallel, whereas the LSTM needs to be processed sequentially, *i.e.*, subsequent steps depend on the previous ones.

Furthermore, we used the BIDMC dataset with the same hyper-parameters to show the generalizability of our proposed method. The input size of the model was set to 256 due to the higher sampling rate of 125 Hz of the BIDMC dataset. Also, the total length of recordings in BIDMC are only 26% of the total length of the recordings in UQVSD, which means that the model was trained with fewer data. Table IV shows that the performance to reconstruct the ECG in BIDMC is similar to that for UQVSD. The higher R peak failure rate with lower MLE and MME is probably due to the noisy data from the critically-ill patients in the BIDMC dataset. This demonstrates that our work can not only perform well with clean normal sinus rhythm but also with more noisy and abnormal signals.

[8]The term "vanilla" means the original architecture without any modification.

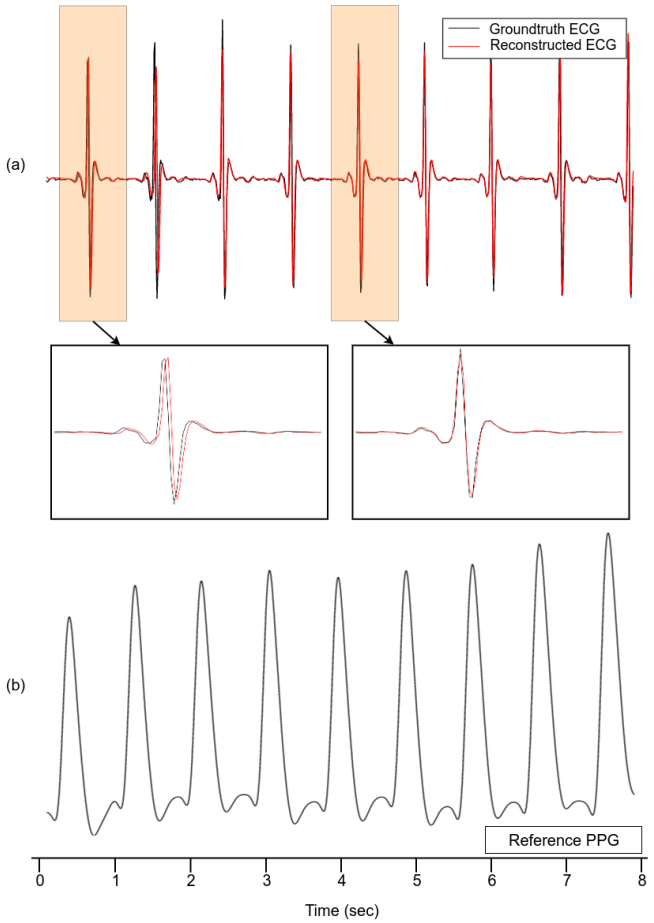


Fig. 3. The reconstructed ECG (red line) and the groundtruth ECG (black line) waveform of 8 seconds in the UQVSD dataset.

D. Qualitative Results

Fig. 3 illustrates 8-second segments of the reconstructed results in the testing set, which shows that the proposed model reconstructs the ECG (Fig. 3(a)) from the reference PPG (Fig. 3(b)) with only a small difference. To investigate different results of the magnitude of R peaks, we estimated four different models (see Fig. 4). Specifically, we stacked the groundtruth ECG and the reconstructed ECG aligned by R peaks at the bottom right of each sub-figure. Our proposed model shows the best performance for the R peak reconstruction (Fig. 4(a)), while our model without the QRS complex-enhanced loss shows a larger lose in the magnitude as shown in Fig. 4(b). In addition, the Vanilla CNN-based

TABLE II
THE PERFORMANCE COMPARISON FOR THE UQVID DATASET.

Model	$L1_{QRS}$	$L1_{nQRS}$	MLE	MME	R_{FAIL}	NMAE	NRMSE
Vanilla LSTM	0.3421	0.0688	1.927	0.343	6.66%	0.2741	0.1081
Ours	**0.3390**	**0.0524**	**1.611**	**0.118**	**3.67%**	**0.2382**	**0.1070**

TABLE III
COMPARISONS OF RUNNING TIME WITH DIFFERENT BATCH SIZES.

Model	#Parameters	1	10	100
Vanilla LSTM	10.6M	1.513ms	1.559ms	2.502ms
Ours	10.8M	1.514ms	1.546ms	2.054ms

TABLE IV
THE PERFORMANCE COMPARISON ON THE UQVID AND BIDMC DATASETS.

Dataset	$L1_{QRS}$	$L1_{nQRS}$	MLE	MME	R_{FAIL}
UQVSD	0.3390	0.0688	1.611	0.118	3.67%
BIDMC	0.3202	0.0386	1.667	0.108	4.37%

TABLE V
PERFORMANCES EVALUATION WITH DIFFERENT σ OF THE UQVSD DATASET. NOTE THAT THE VALUE IN THE BRACKET REPRESENTS THE VARIANCE.

	MLE	MME	R_{FAIL}
1	**1.641 (5.121)**	**0.111 (0.015)**	**3.82%**
3	1.665 (5.365)	0.172 (0.022)	4.27%
5	1.683 (5.417)	0.182 (0.023)	4.39%
7	1.656 (5.349)	0.190 (0.024)	4.31%

transform network performs competitively compared to our model in Fig. 4(c), but the overall evaluation is worse than that of our model (see the next section for a discussion). Moreover, Fig. 4(d) shows that the Vanilla LSTM has the worst performance since the magnitude of the R peaks is much smaller than that of the groundtruth. In summary, it is easy for other models to reconstruct the location of R peaks accurately, but it is hard to calculate the magnitude of the R peaks as correctly as the proposed model does.

E. Sensitivity Analysis and Ablation Study

Table V shows the performance of the proposed model with different ranging from 1 to 7, and it is clear that = 1 yields the best R peak failure rate, MLE and MME. When is too large, the performance worsens since a large takes a long period into consideration. Moreover, Table VI lists the performance indicators of the proposed model with different (0; 0.1; 0.5; 1; 10). The results indicate that = 0.5 is the best

TABLE VI
PERFORMANCES EVALUATION WITH DIFFERENT β OF THE UQVSD DATASET. NOTE THAT THE VALUE IN THE BRACKET REPRESENTS THE VARIANCE.

	MLE	MME	R_{FAIL}
0	1.690 (5.572)	0.209 (0.027)	5.05%
0.1	1.632 (5.086)	0.109 (0.014)	3.84%
0.5	**1.616 (5.074)**	**0.109 (0.014)**	**3.75%**
1	1.633 (5.224)	0.110 (0.015)	4.02%
10	1.630 (5.193)	0.111 (0.014)	4.05%

for the UQVID dataset, because it achieves the best MME, MLE and also the lowest R peak failure rate. It is worth noting that is used to penalize the loss in the region of the QRS complex, and = 0 is the L1 loss and shows the worst performance in terms of the R peak failure rate. According to the sensitivity test, we recommend setting the default values of and to 1 and 0.5, respectively. After this, we investigated the performance of different modules: sequence transformer network, attention network, and QRS complex-enhanced loss. Table VII shows the 8 different combinations of these three modules. The findings indicate that the sequence transformer network improves the model with the attention network and the QRS complex-enhanced loss. If we disable the sequence transformer network, the MLE and the variance of MLE and MME increase, which leads to a high R peak failure rate (rows 1, 5). In contrast, if we enable it in the Vanilla model, only MLE shows improvement and MME gets worse (rows 3, 4). Moreover, the attention network is more important for the R peak location than the R peak magnitude in the full model (rows 1, 6), and it helps both MLE and MME compared to the vanilla model (rows 4, 7). The attention network adds the ability to the transform network for choosing important samples to reconstruct the ECG.

Furthermore, Table VII shows that QRS complex-enhanced loss plays an essential role in R peak magnitude, as it helps every model to get a significantly low MME. Nonetheless, the performance of QRS complex-enhanced loss is not good enough for better R peak detection since both location and magnitude are indispensable. In other words, combining all modules is necessary to get low MLE and MME, as well as a lower R peak failure rate. In summary, the sequence transformer network and attention network are essential to reconstruct R peaks at an accurate location, while the QRS complex-enhanced loss is critical for the magnitude of R peaks. All of them have an interaction effect to improve the overall performance. Finally, from the evaluation of the impact of data augmentation, Table VIII shows that the model with data augmentation significantly improves for R peak location, because the data augmentation offers opportunities to learn from shifted samples with R peaks at the boundaries. Therefore, data augmentation plays an important role in stabilizing the training.

To show the generalizability of the proposed QRS complex-enhanced loss, we apply it to Vanilla LSTM. Table IX shows the results of different models with the proposed QRS complex-enhanced loss, which manifest that the MLE, MME and R_{FAIL} in the model with QRS complex-enhanced loss are significantly better than that in the model without QRS complex-enhanced loss.

TABLE VII
ABLATION STUDY IN TEST SET OF THE UQVID DATASET. NOTE THAT THE VALUE IN THE BRACKET REPRESENTS THE VARIANCE.

STN	Attention	L_{QRS}	MLE	relative MLE	MME	relative MME	R_{FAIL}
X	X	X	**1.611 (5.059)**	+0.00%	0.118 (0.016)	+3.51%	**3.67%**
X	X		1.620 (5.149)	+0.56%	0.208 (0.028)	+82.46%	4.49%
X			1.648 (5.325)	+2.30%	0.221 (0.030)	+93.86%	4.59%
			1.690 (5.572)	+4.90%	0.209 (0.027)	+83.33%	5.05%
	X	X	1.624 (5.228)	+0.81%	**0.114 (0.016)**	+0.00%	3.75%
X		X	1.631 (5.111)	+1.24%	0.117 (0.016)	+2.63%	4.04%
	X		1.656 (5.522)	+2.79%	0.210 (0.028)	+84.21%	4.96%
		X	1.657 (5.214)	+2.86%	0.116 (0.016)	+1.75%	3.91%

TABLE VIII
IMPROVEMENT BY DATA AUGMENTATION OF THE UQVSD DATASET. NOTE THAT THE VALUE IN THE BRACKET REPRESENTS THE VARIANCE.

Model	MLE	MME	R_{FAIL}
Ours	**1.611 (5.059)**	**0.118 (0.016)**	**3.67%**
Ours w/o DA	1.828 (5.978)	0.121 (0.022)	4.45%

TABLE IX
EVALUATION OF COMPATIBILITY OF QRS COMPLEX-ENHANCED LOSS TO VANILLA LSTM OF UQVSD DATASET.

Model	MLE	MME	R_{FAIL}
Vanilla LSTM	1.927 (6.544)	0.343 (0.035)	6.66%
w/ L_{QRS}	**1.745 (5.474)**	**0.133 (0.015)**	**3.91%**

F. Possible Application

Our model is valuable in its potential for reconstructing ECG because it is equipped with the direct transformation from PPG to ECG without the need to measure both PPG and ECG simultaneously, it is valuable to see the potential of the reconstructed ECG from our model. One important contribution of this study is that we preserve the temporal information between PPG and ECG, so we can compute the pulse transit time (PPT) as shown in Fig. 5, which is important for blood pressure estimation [33] [34], respiratory effort [35], heart rate [36] and other vital signs. Our proposed method has the capability to retrieve the PPT from the input PPG and the reconstructed ECG by measuring the time difference between their peaks. Because the test samples are serialized by using a sliding window, we calculated the PPT by considering two consecutive test cases for the possible R peak that occurred in the following test case. After that, we filtered out the PPT values at the range [480ms, 800ms] according to previous research [37], which pointed out that the time from the R peak of ECG to the peak of PPG is 80% of the time of the R-R interval of the ECG, and the normal resting heart rate of adults is within the range of 60 and 100 beats per minute of heart rate. From this we observed that, the detection of R peaks contained 323 failed PPT among 5483 R peaks in total.[9] As can be seen in Fig. 6, we used the Pearson correlation coefficient of the PPT from the groundtruth and reconstructed signals to evaluate our model. The results indicate that most of the PPT fell between 550ms and 750ms, while the correlation coefficient was equal to 0.844, which indicates that the PPT from the reconstructed ECG is similar to the groundtruth. The results also show that the proposed sequence transformer network and attention network successfully reconstructed the ECG, and that the QRS complex-enhanced loss is important for R peak magnitude. This further demonstrates the potential of our model to exploit further measurements for clinical study. In other words, we only need to use the PPG to reconstruct ECG and then extract meaningful information from the reconstructed ECG, which has been well-studied.

[9]The 5.89% failure rate is possibly due to the instability of the peak detection algorithm, which can be further improved in the future.

V. CONCLUSION AND FUTURE WORK

In this paper, we proposed the first end-to-end learning framework to directly transform PPG into ECG, thus leveraging the advantages of both sides: PPG is economical and easily accessible while ECG has well-studied base. The proposed model consists of a sequence transformer network and an attention network that help the main transform network to accurately reconstruct ECG from only PPG. Also, a new QRS complex-enhanced loss is introduced to make the model more robust. We tested the proposed approach using two open datasets (the University of Queensland Vital Signs Dataset and the Beth Israel Deaconess Medical Centre Dataset). The experiment results confirmed the effectiveness and generalizability of the proposed deep learning approach, which successfully reconstructs ECG from only PPG as input and has the ability to make inferences in real-time. Additional findings show the importance of each module in our model and provide a direction to retrieve a better quality for the QRS complex. In conclusion, this study suggests that deep learning can be a potential and viable approach for the reconstruction of vital signs from other easily accessible signals. In the future, we will address the model's current limitation of needing personally paired PPG and ECG data for training by utilizing a new model architecture and few-shot learning skills.

VI. ACKNOWLEDGEMENT

This work was supported in part by the Ministry of Science and Technology of Taiwan under Grants MOST-108-2221-E-009-088, MOST-108-2622-E-009-026-CC2, MOST-109-2622-8-009-018-TE1, and MOST-108-2823-8-009-002. We are grateful to the National Center for High-performance Computing for computer time and facilities.

REFERENCES

[1] P. K. et al., "Recommendations for the standardization and interpretation of the electrocardiogram: Part i: The electrocardiogram and its technology,," *Journal of the American College of Cardiology*, vol. 49, pp. 1109–1127, Jan. 2007.

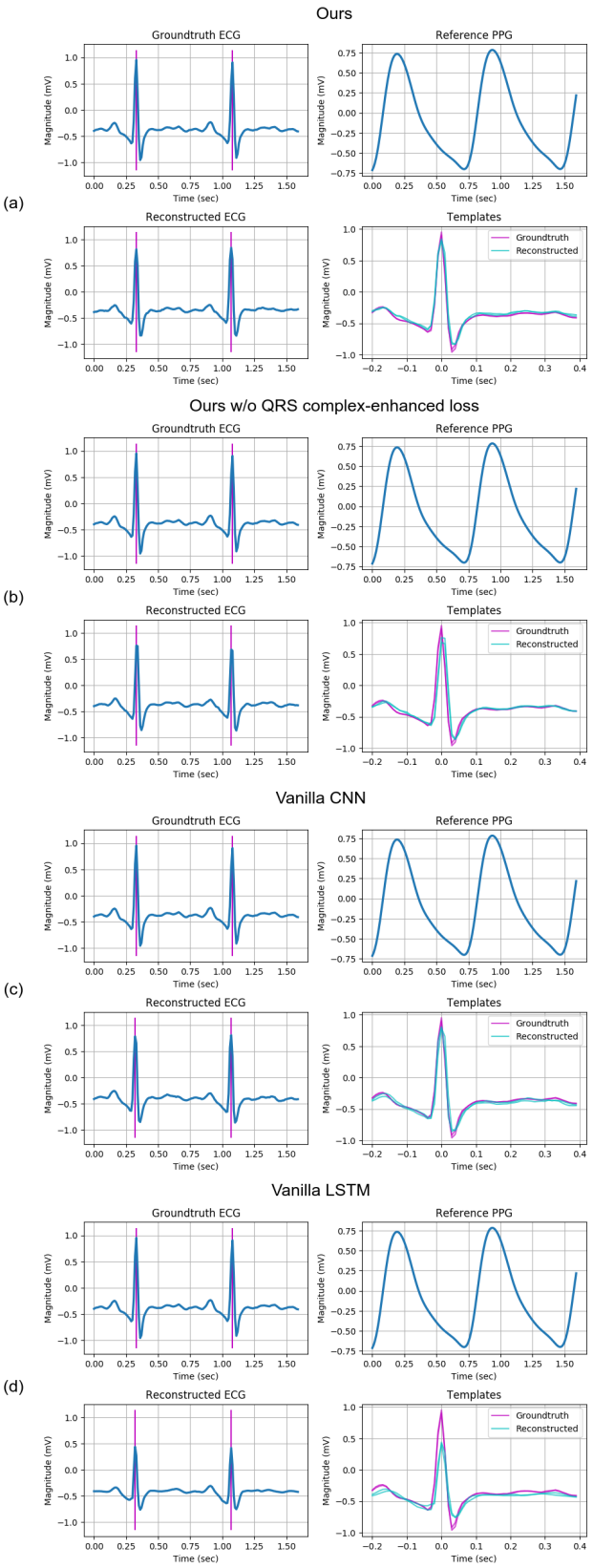


Fig. 4. The analysis of R peaks in the UQVSD dataset.

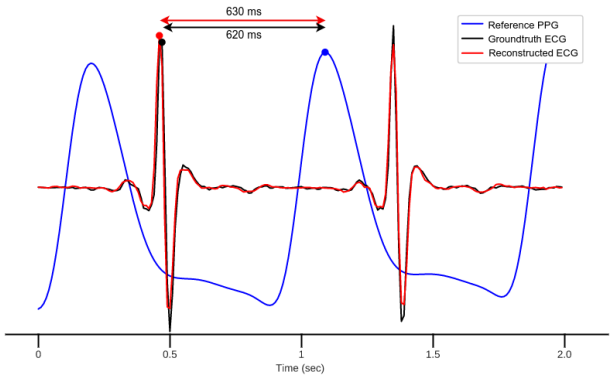


Fig. 5. The pulse transit time from the groundtruth signals and reconstructed signals in the UQVSD dataset.

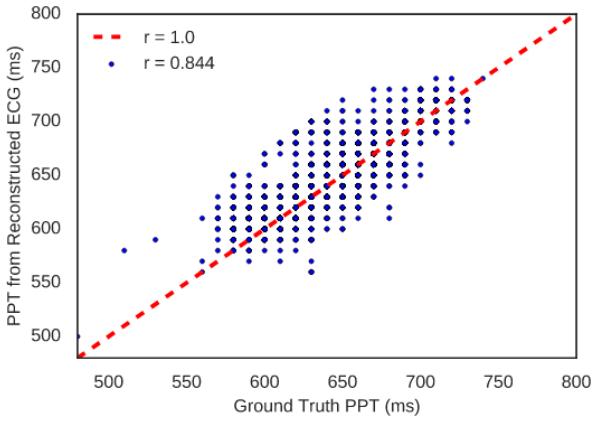


Fig. 6. The scatter plot illustrates the relationship between pulse transit time calculated from the groundtruth and from the reconstructed ECG in the UQVSD dataset.

[2] A. Reisner, P. A. Shaltis, D. McCombie, and H. H. Asada, "Utility of the photoplethysmogram in circulatory monitoring," *Anesthesiology: The Journal of the American Society of Anesthesiologists*, vol. 108, pp. 950–958, May 2008.

[3] M. Shafique, P. A. Kyriacou, and S. K. Pal, "Investigation of photoplethysmographic signals and blood oxygen saturation values on healthy volunteers during cuff-induced hypoperfusion using a multimode ppg/spo2 sensor," *Medical & Biological Engineering & Computing*, vol. 50, pp. 575–583, Jun 2012.

[4] Shelley and K. H., "Photoplethysmography: Beyond the calculation of arterial oxygen saturation and heart rate," *Anesthesia & Analgesia*, vol. 105, Dec 2007.

[5] X. F. Teng and Y. T. Zhang, "Continuous and noninvasive estimation of arterial blood pressure using a photoplethysmographic approach," in *IEEE International Conference on Engineering in Medicine and Biology Society (EMBC)*, Sep 2003.

[6] L. Wang, E. Pickwell-MacPherson, Y. P. Liang, and Y. T. Zhang, "Non-invasive cardiac output estimation using a novel photoplethysmogram index," in *IEEE International Conference on Engineering in Medicine and Biology Society (EMBC)*, Sep 2009.

[7] L. Nilsson, A. Johansson, and S. Kalman, "Monitoring of respiratory rate in postoperative care using a new photoplethysmographic technique," *Journal of Clinical Monitoring and Computing*, vol. 16, pp. 309–315, May 2000.

[8] D. Castaneda, A. Esparza, M. Ghamari, C. Soltanpur, and H. Nazeran, "A review on wearable photoplethysmography sensors and their potential future applications in health care," *International Journal of Biosensors & Bioelectronics*, vol. 4, pp. 195–202, Aug 2018.

[9] M. Chen, Z. Xu, K. Q. Weinberger, and F. Sha, "Marginalized denoising autoencoders for domain adaptation," in *International Conference on Machine Learning (ICML)*, Jun 2012.

[10] M. Sakurada and T. Yairi, "Anomaly detection using autoencoders with nonlinear dimensionality reduction," in *Workshop on Machine Learning for Sensory Data Analysis (MLSDA)*, Dec 2014.

[11] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. van der Laak, B. van Ginneken, and C. I. Sánchez,

"A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, p. 60–88, Dec 2017.
[12] S. M. Erfani, S. Rajasegarar, S. Karunasekera, and C. Leckie, "High-dimensional and large-scale anomaly detection using a linear one-class svm with deep learning," *Pattern Recognition*, vol. 58, pp. 121–134, Oct 2016.
[13] Y. A. Bhagat, P. Vcrdon, S. Avuthu, D. Parsons, M. Sussman, G. Wable, and R. Hugeneck, "Like kleenex for wearables: A soft, strong and disposable ecg monitoring system," in *IEEE Biomedical Circuits and Systems Conference (BioCAS)*, Oct 2018.
[14] E. Spanò, S. Di Pascoli, and G. Iannaccone, "Low-power wearable ecg monitoring system for multiple-patient remote monitoring," *IEEE Sensors Journal*, vol. 16, pp. 5452–5462, Jul 2016.
[15] S. Majumder, L. Chen, O. Marinov, C. Chen, T. Mondal, and M. J. Deen, "Noncontact wearable wireless ecg systems for long-term monitoring," *IEEE Reviews in Biomedical Engineering*, vol. 11, pp. 306–321, May 2018.
[16] H. T. Haverkamp, S. O. Fosse, and P. Schuster, "Accuracy and usability of single-lead ecg from smartphones - a clinical study," *Indian Pacing and Electrophysiology Journal*, vol. 19, no. 4, pp. 145–149, 2019.
[17] Z. Cai, J. Li, X. Zhang, Q. Shen, A. Murray, and C. Liu, "How accurate are ecg parameters from wearable single-lead ecg system for 24-hours monitoring," in *Computing in Cardiology (CinC)*, Sep 2019.
[18] R. Banerjee, A. Sinha, A. D. Choudhury, and A. Visvanathan, "Photoecg: Photoplethysmographyto estimate ecg parameters," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2014.
[19] Q. Zhu, X. Tian, C.-W. Wong, and M. Wu, "Ecg reconstruction via ppg: A pilot study," in *IEEE EMBS International Conference on Biomedical Health Informatics (BHI)*, May 2019.
[20] L. D., G. M., and Jenkins, "The university of queensland vital signs dataset: Development of an accessible repository of anesthesia patient monitoring data for research," *Anesthesia & Analgesia*, vol. 114, pp. 584–589, Dec. 2012.
[21] M. A. F. Pimentel, A. E. W. Johnson, P. H. Charlton, D. Birrenkott, P. J. Watkinson, L. Tarassenko, and D. A. Clifton, "Toward a robust estimation of respiratory rate from pulse oximeters," *IEEE Transactions on Biomedical Engineering*, vol. 64, pp. 1914–1923, Aug 2017.
[22] J. Oh, J. Wang, and J. Wiens, "Learning to exploit invariances in clinical time-series data using sequence transformer networks," in *Machine Learning for Healthcare Conference (MLHC)*, Aug 2018.
[23] B. Wang, K. Liu, and J. Zhao, "Inner attention based recurrent neural networks for answer selection," in *Annual Meeting of the Association for Computational Linguistics (ACL)*, Aug 2016.
[24] Pascual, Santiago, Bonafonte, Antonio, and J. Serrà, "Segan: Speech enhancement generative adversarial network," in *Annual Conference of the International Speech Communication Association (INTERSPEECH)*, Aug 2017.
[25] S. Iizuka, E. Simo-Serra, and H. Ishikawa, "Globally and locally consistent image completion," in *Special Interest Group on Computer Graphics and Interactive Techniques (SIGGRAPH)*, Jul 2017.
[26] R. Silipo, M. Gori, A. Taddei, M. Varanini, and C. Marchesi, "Classification of arrhythmic events in ambulatory electrocardiogram, using artificial neural networks," *Computers and Biomedical Research*, vol. 28, pp. 305–318, Aug 1995.
[27] P. Hamilton, "Open source ecg analysis," in *IEEE Conference on Computers in Cardiology (CinC)*, Sep 2002.
[28] C. Carreiras, A. P. Alves, A. Lourenço, F. Canento, H. Silva, A. Fred, *et al.*, "BioSPPy: Biosignal processing in Python," 2015.
[29] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *International Conference on Learning Representations (ICLR)*, Dec 2014.
[30] N. Darouian, K. Narayanan, A. L. Aro, K. Reinier, A. Uy-Evanado, C. Teodorescu, K. Gunson, J. Jui, and S. S. Chugh, "Delayed intrinsicoid deflection of the qrs complex is associated with sudden cardiac arrest," *Heart Rhythm*, vol. 13, pp. 927–932, Apr 2016.
[31] C. Meyer, J. F. Gavela, and M. Harris, "Combining algorithms in automatic detection of qrs complexes in ecg signals," *IEEE Transactions on Information Technology in Biomedicine*, vol. 10, pp. 468–475, Jul 2006.
[32] G. K. Anumanchipalli, J. Chartier, and E. F. Chang, "Speech synthesis from neural decoding of spoken sentences," *Nature*, vol. 568, pp. 493–498, Apr 2019.
[33] L. A. Geddes, M. H. Voelz, C. F. Babbs, J. D. Bourland, and W. A. Tacker, "Pulse transit time as an indicator of arterial blood pressure," *Psychophysiology*, vol. 18, pp. 71–74, Jan 1981.
[34] R. A. Payne, C. N. Symeonides, D. J. Webb, and S. R. J. Maxwell, "Pulse transit time measured from the ecg: an unreliable marker of beat-to-beat blood pressure," *Journal of Applied Physiology*, vol. 100, pp. 136–141, Jan 2006.
[35] R. Smith, J. Argod, J. Pépin, and P. Lévy, "Pulse transit time: An appraisal of potential clinical applications," *Thorax*, vol. 54, pp. 452–457, Jun 1999.
[36] M. J. Drinnan, J. Allen, and A. Murray, "Relation between heart rate and pulse transit time during paced respiration," *Physiological Measurement*, vol. 22, pp. 425–432, Jul 2001.
[37] A. Vest, Q. Li, C. Liu, S. Nemati, G. Da Poian, A. Shah, and G. Clifford, "An open source benchmarked toolbox for cardiovascular waveform and interval analysis," *Physiological Measurement*, vol. 39, p. 105004, Oct 2018.

Hong-Yu Chiu received the B.S. degrees from the Department of Life Sciences, National Yang-Ming University (NYMU), Taipei, Taiwan, R.O.C. and the Department of Electrical and Computer Engineering, National Chiao Tung University (NCTU), Hsinchu, Taiwan, R.O.C., in 2019. His research interests include artificial intelligence, deep learning, and computer vision.

Hong-Han Shuai received the B.S. degree from the Department of Electrical Engineering, National Taiwan University (NTU), Taipei, Taiwan, R.O.C., in 2007, the M.S. degree in computer science from NTU in 2009, and the Ph.D. degree from Graduate Institute of Communication Engineering, NTU, in 2015. He is now an assistant professor in NCTU. His research interests are in the area of multimedia processing, machine learning, social network analysis, and data mining. His works have appeared in top-tier conferences such as MM, CVPR, AAAI, KDD, WWW, ICDM, CIKM and VLDB, and top-tier journals such as TKDE, TMM and JIOT.

Paul C.-P. Chao received his Ph.D. degree from Michigan State University, USA, and then with Chrysler Corp, Detroit, USA before joined National Chiao Tung University (NCTU), Taiwan. He is currently University Distinguished Professor of the electrical engineering department at NCTU, and Distinguished Lecturer for IEEE Sensors Council, 2018 – 2020. His research interests focus on sensors, actuators and their interface circuitry. Dr. Chao has published more than 300 peer-reviewed papers (books, journal papers, conferences, reports) and 38 patents. Dr. Chao was the recipient of the 1999 Arch T. Colwell Merit Award from Society of Automotive Engineering, Detroit, USA; The 2019 Technical Achievement Award, IEEE Sensors Council; The 2017 Presidential Outstanding Professor of Engineering in Nation (Taiwan) (awarded by the president of the nation in the Presidential House of Taiwan, ROC); Two 2017 Future Technology Awards (Taiwan Oscar Invention Award) from Ministry of Science and Technology (MOST), Taiwan Government; The 2018 Outstanding Professor of Electrical Engineering in Nation (Taiwan), National Association of Electrical Engineering, Taiwan; the Outstanding research Award of MOST, 2018. Dr. Chao has served as University Associate Vice Presidents of NCTU for academic affairs (2009-2010) and research and development (2015). Dr. Chao is currently Topical Editors of IEEE Sensors Journal and IEEE IoT Journal. Dr. Chao also received the award of the 2017 Best Topical Editor, Runner up, IEEE Sensors Journal. He is a senior member of IEEE and ASME Fellow.