

Article

Data Independent Acquisition Based Bi-Directional Deep Networks for Biometric ECG Authentication

Htet Myet Lynn ¹ , Pankoo Kim ¹ and Sung Bum Pan ^{2,*} 

¹ Department of Computer Engineering, Chosun University, Gwangju 61452, Korea; htetmyet@chosun.kr (H.M.L.); pkkim@chosun.ac.kr (P.K.)

² Department of Electronics Engineering, Chosun University, Gwangju 61452, Korea

* Correspondence: sbpan@chosun.ac.kr

Abstract: In this report, the study of non-fiducial based approaches for Electrocardiogram (ECG) biometric authentication is examined, and several excessive techniques are proposed to perform comparative experiments for evaluating the best possible approach for all the classification tasks. Non-fiducial methods are designed to extract the discriminative information of a signal without annotating fiducial points. However, this process requires peak detection to identify a heartbeat signal. Based on recent studies that usually rely on heartbeat segmentation, QRS detection is required, and the process can be complicated for ECG signals for which the QRS complex is absent. Thus, many studies only conduct biometric authentication tasks on ECG signals with QRS complexes, and are hindered by similar limitations. To overcome this issue, we proposed a data-independent acquisition method to facilitate highly generalizable signal processing and feature learning processes. This is achieved by enhancing random segmentation to avoid complicated fiducial feature extraction, along with auto-correlation to eliminate the phase difference due to random segmentation. Subsequently, a bidirectional recurrent neural network (RNN) with long short-term memory (LSTM) deep networks is utilized to automatically learn the features associated with the signal and to perform an authentication task. The experimental results suggest that the proposed data-independent approach using a BLSTM network achieves a relatively high classification accuracy for every dataset relative to the compared techniques. Moreover, it exhibited a significantly higher accuracy rate in experiments using ECG signals without the QRS complex. The results also revealed that data-dependent methods can only perform well for specified data types and amendments of data variations, whereas the presented approach can also be considered for generalization to other quasi-periodical biometric signal-based classification tasks in future studies.

Keywords: ECG; authentication; RNN; LSTM; deep networks



Citation: Lynn, H.M.; Kim, P.; Pan, S.B. Data Independent Acquisition Based Bi-Directional Deep Networks for Biometric ECG Authentication. *Appl. Sci.* **2021**, *11*, 1125. <https://doi.org/10.3390/app11031125>

Received: 10 November 2020

Accepted: 20 January 2021

Published: 26 January 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

At present, we are experiencing emerging digitization in most aspects of our lives. Day-to-day, use online applications and services such as mobile banking, social networking, online stock exchange, and trading or email services, with little apprehension about storing personal confidential information on our devices or client servers. In the digital era, sadly, we are also faced with new attacks and exploits, as well as unauthorized access to sensitive information and devices by malicious viruses or hosts. It is incredible that large populations of users still rely on numerous types or particular sets of passwords, which have been used for authorized access since the earliest era of computing. In recent years, there has been a shift in attention towards biometric security systems. These security applications facilitate the identification of an individual using their distinct biological characteristics instead of a set of numerical or alphabetical passwords. The most widespread techniques use fingerprint, iris, and facial recognition approaches, and are normally found in smart devices [1]. Regarding biometrics in mobile devices, the benefits are apparent. There is an added degree of security in relying, at least in part, on an extremity (e.g., finger) that only the user

possesses. For instance, with a standard password, together with a personal fingerprint, the sense of security is enhanced. Beyond smartphone security, the use of features such as touch IDs improves the convenience of interacting with your device. With applications based on such security systems, instead of manually entering your payment information, users simply have to swipe a finger across their device. It is simple and saves time. Using biometric authentication for security purposes also improves the security of information, processes, and establishments. Some organizations have implemented biometric scanning as a modern method of "punching in" to work. This assures that all employees are honest in terms of the hours they have worked. In turn, this saves the organization money. However, there are still difficulties and issues related to fingerprint usability and reliability. Current challenges in ECG biometric classification [2–4] tasks include the extraction of features from the ECG signals to implement a model to learn hidden patterns for accurate generalization, proving the stability of the biometric and protecting against attacks. In this report, we proposed biometric-based ECG signals for human authentication with non-fiducial techniques.

The ECG signal usually consists of three complexes, namely P, QRS, and T. They are determined based on their corresponding complexes, also known as fiducial points, which are the peak points of the respective complex. Using this information as distinct features, more informative characteristics including time-domain features such as amplitudes and intervals, are generally used as features for individual signals. However, the shapes of the ECG signal can vary depending on the location of the electrode on the human body during data acquisition. P-QRS-T complexes are not available for every version of the ECG signal. Thus, a non-fiducial approach using machine learning is used to address these problems. By introducing a non-fiducial method based on deep learning techniques, the fiducial extraction can be neglected in the pre-processing phase. In addition, unlike previous studies that relied on specified data for classification, a data-independent acquisition technique is also extended to form a model that can handle various types of ECG data input. To enable a data-independent and highly generalizable signal processing and feature learning process, a random segmentation based a wavelet domain multiresolution bi-directional LSTM network is proposed. Specifically, it allows for blindly selecting a physiological signal segment for classification purpose, avoiding the complicated signal fiducial characteristics extraction process which does not rely on QRS peak detection. It also removes phase difference among random-chosen signal segments by the auto-correlation approach. Thus, ECG data with either QRS peak or without QRS peak can be processed for data acquisition task in upstreaming phase without any hesitation. However, the scope of our research was focused on non-fiducial approach, rather than fiducial techniques, our proposed models are examined only on non-fiducial approach. The major contributions of this research are as follows:

1. Applied random segmentation and auto-correlation for various types of ECG data input independently, and to produce a reasonable quantity of training data from a raw signal [5].
2. Proposed and compared the performance of generalization by designing 1D-CNN networks, bidirectional RNNs on both Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) cells.

The report consists of four components. The first is a brief review of related works and the methodology. In the second part, the authors present their research methodology and the detailed explanation of pre-processing the data with insightful information on the proposed method. In the third part, the research results and comparisons with other techniques are presented, and in the final, section, the main conclusions and concluding remarks are presented.

2. Related Work

2.1. Fiducial Methods

ECG-based authentication applications are typically based on two approaches, namely, fiducial and non-fiducial methods. The fiducial approach utilizes a feature extraction process, wherein the points of interest are within a heartbeat wave. Usually, the heartbeat wave consists of a P-QRS-T complex, which represents the activity of a beating human heart, acquired using electrodes placed on different parts of the body. These complexes are then used to extract latency and amplitude features [4,6]. Such approaches generally rely on robust heartbeat segmentation and fiducial peak point detection [7–10]. Thus, manual feature engineering efforts are necessary to capture a heartbeat from an original ECG signal. Moreover, precise estimation is important to annotate the peak locations of the P-QRS-T complex of a signal. After this fiducial information is detected, the amplitude and time-interval between corresponding points are measured. The entire process of identifying peak data points of the QRS complex and calculating the time-interval features are considered as the pre-processing component or the signal processing phase in ECG-based biometric authentication applications, as shown in Figure 1.

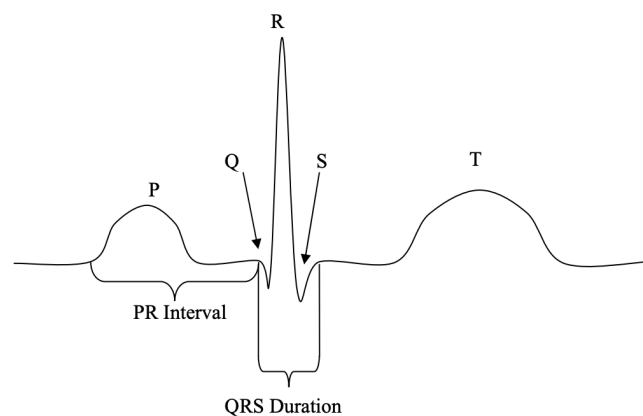


Figure 1. Fiducial points of P-QRS-T complex of ECG signal.

2.2. Non-Fiducial Methods

Numerous recent studies have been proposed in the field of biometric signal-based security systems by exploiting deep learning methods [11–18]. Q. Zhang et al. [19] proposed a multi-resolution network based on 1D-CNN for ECG human identification applications for smart devices. The method was extended to transform raw input signals into multiple versions of wavelets to improve the context representations of the signal. However, the auto-correlation of segmented windows and transformation of the wavelet are necessary. X. Zhang [20] proposed models using RNN networks in various types of cell units in hidden layers. The results suggested that the use of both LSTM and GRU gates did not significantly affect performance in terms of classification accuracy. M. Al Rahhal et al. [21] proposed a method that implements stacked denoising auto-encoders with sparsity constraints, and a softmax layer was applied on top of the hidden representation layer as a deep neural network. M. Zihlmann et al. [22] proposed two models based on deep neural networks, CNN, and a hybrid approach for combining the CNN and the RNN network using an LSTM cell unit. A similar approach was proposed by Warrick and Homosi [18], which automatically learned the hidden characteristics of a signal and identified cardiac arrhythmias in an ECG signal using CNN and LSTM techniques.

2.2.1. Convolutional Neural Networks (CNN)

Convolutional neural networks are a category of deep neural networks that have proven to be effective in areas such as image recognition and related classification tasks. CNNs have been successful in identifying faces [23], objects, and traffic signs. It uses convolutional layers to filter input data to obtain useful information, and a nonlinear

activation function is applied to the results of the convolutional operation. The convolution operation includes the combination of inputs with the kernel, also known as filters, to form a transformed feature map. Then, a fully connected layer is used after the pooling process for classification. The kernel filters complete features extracted by sliding from top to bottom and from left to right in the original matrix. A convolutional neural network is also known as a type of multi-layer neural network that improves the error backpropagation network. CNNs are good at classifying images, especially large images. CNN was first proposed by Y. Lecun and used for handwritten character recognition [24], see in Figure 2.

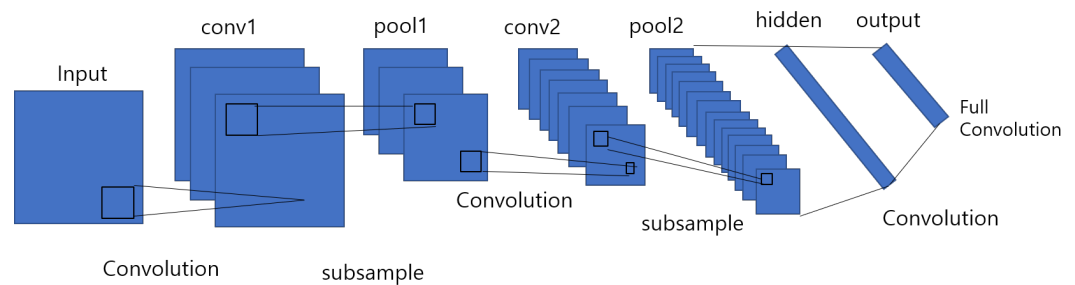


Figure 2. A sample of 7-layer CNN model classified digits for digitized pixel greyscale input images in image processing.

The convolutional neural network (CNN) technique has two components: a feature identifier and a fully connected layer. The feature identifier is performed using convolutional layers and pooling layers, wherein the features are learned automatically. In ECG-based classification problems, the fully connected component performs signal classification using the features learned from the feature's identifier component.

2.2.2. Recurrent Neural Networks (RNN)

The RNN has is highly preferred method [25,26], especially for sequential data. Every node at a time step consists of an input from the previous node, and it proceeds using a feedback loop. In RNN, each node generates a current hidden state and its output is obtained by using the given input and previous hidden state as follows:

$$h_t = f(W_h h_{t-1} + V_h x_t + b_h) \quad (1)$$

$$o_t = f(W_o h_t + b_o) \quad (2)$$

where indicates the hidden block of each time step t . W and V are the weights for the hidden layers, b denotes the bias for the hidden and output states, and f denotes the activation function applied on each node throughout the network. o_t is the output vector at time step t which predict the next output of the sequence from the network, and is a softmax function of hidden representation h and weights W associated with it along with the bias b .

2.2.3. Long Short-Term Memory (LSTM)

Long short-term memory is a type of RNN model designed to prevent the output of a neural network from either exploding or decaying (long-term dependency) as it passes through the feedback loops for a given input [27]. Such feedback loops in RNNs allow the network to be better at pattern recognition compared to other neural networks. Due to their ability to learn long-term dependency, LSTMs are applicable to numerous long sequence learning problems such as language modeling and machine translation, and many other related tasks. LSTM models are designed by applying memory cells with several gates in a hidden layer, as shown in Figure 3a. The hidden layer blocks with the LSTM cell unit, and the three functions of gate controllers are formulated as follows:

- Forget gate f_t decides which part of long-term state c_t should be omitted.
- Input gate i_t controls which part should be added to long-term state.
- Output gate g_t determines which part of c_t should be read and outputs to h_t and o_t .

The following equations facilitate the calculation of the long-term and short-term states of the cell and the output of each layer in the time step.

$$f_t = \sigma(W_{x,f}^T x_t + W_{h,f}^T h_{t-1} + b_f) \tag{3}$$

$$i_t = \sigma(W_{x,i}^T x_t + W_{h,i}^T h_{t-1} + b_i) \tag{4}$$

$$o_t = \sigma(W_{x,o}^T x_t + W_{h,o}^T h_{t-1} + b_o) \tag{5}$$

$$g_t = \tanh(W_{x,g}^T x_t + W_{h,g}^T h_{t-1} + b_g) \tag{6}$$

$$c_t = f_t \otimes c_{t-1} + i_t \otimes \hat{c}_t \tag{7}$$

$$o_t, h_t = g_t \otimes \tanh(c_t) \tag{8}$$

where $W_{x,f}, W_{x,i}, W_{x,o}, W_{x,g}$ denote the weight parameters for the connected input vector, $W_{h,f}, W_{h,i}, W_{h,o}, W_{h,g}$ denote the weight parameters of the short-term state of the previous time step, and $b_f, b_i, b_o,$ and b_g are the biases.

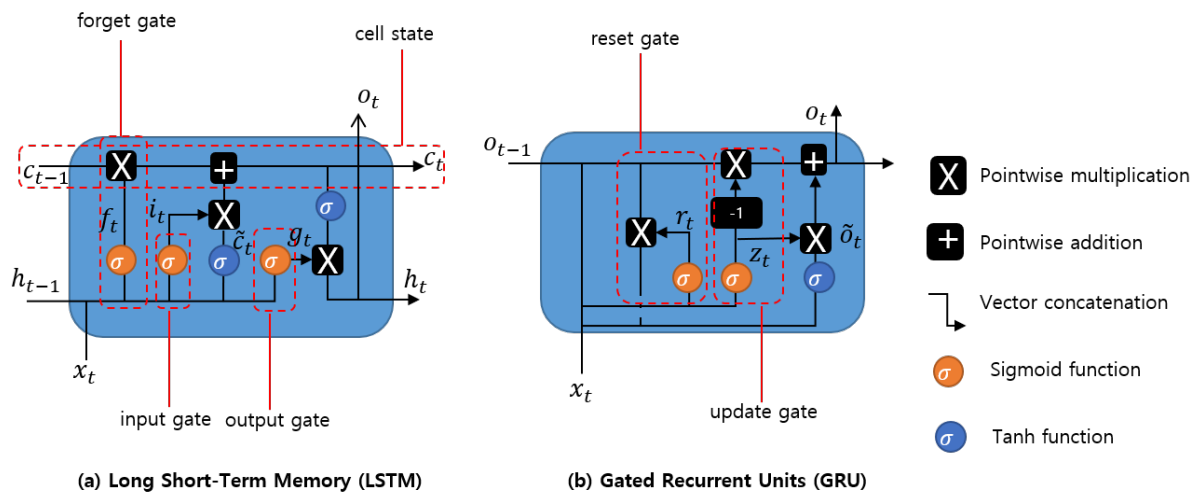


Figure 3. Different cell units and their respective operations of hidden layers in Recurrent Neural Networks; Long ShortTerm Memory, Gated Recurrent Unit.

2.2.4. Gated Recurrent Unit(GRU)

Generally, both LSTM and GRU cell units are applied based on the intuition of avoiding the vanishing gradient problem in deep neural networks. The GRU is modified with an update gate to decide whether to pass the previously hidden layer output to the next cell. The forged gate is implemented for additional mathematical operation with a new set of weights. Intuitively, the reset gate decides to combine the new input with the previous memory. The update gate determines which part of the previous memory information should be passed on to the network to calculate the new state. The insight information and structure is shown in Figure 3b, and the formulations for each gate and their outputs are calculated as follows:

$$r_t = \sigma(W_{x,f}^T x_t + W_{h,f}^T h_{t-1} + b_f) \tag{9}$$

$$z_t = \sigma(W_{x,i}^T x_t + W_{h,i}^T h_{t-1} + b_i) \tag{10}$$

$$\hat{o}_t = \tanh(W_{x,\delta}^T x_t + W_{o,\delta}^T (r_t \otimes o_{t-1}) + b_\delta) \tag{11}$$

$$o_t = z_t \otimes o_{t-1} + (1 - z_t) \otimes \hat{o}_t \tag{12}$$

where $W_{x,r}$, $W_{x,z}$, $W_{x,\delta}$ denote the weight matrices for the corresponding connected input vector, $W_{o,r}$, $W_{o,z}$, $W_{o,\delta}$ represent the weight matrices of the previous time step, and b_r , b_z , and b_δ are the biases.

3. Methodology

3.1. Data Argumentation Process

In this research, ECG-ID (ECGID), MIT-BIH Arrhythmia Database (MIT-BIH ECG), STAFF-III, and LT-AF [28–31] for ECG signals with QRS complex dataset, and AFDB [32], and AHA dataset [33] for ECG signals without the QRS complex from PhysioNet [34], have been performed separately for all candidate models. The signal processing phase can also be determined as data pre-processing in ECG authentication applications. Figure 4 illustrates the framework of data preprocessing phase of proposed methodology where the raw ECG signals with QRS complex are conducted with data argumentation process such as data filtering, R-peak detection and heartbeat segmentation in order to obtain a segmented single heartbeat window before they are fed into proposed deep networks. Similarly, ECG signals without QRS complex are conducted with extended data independent acquisition process where the signals are randomly segmented without having peak detection, and performed auto-correlation technique to transform two dimensional data to be fed into downstreaming process. In data argumentation process, it mainly consists of three core operations, i.e., detrending, noise removal (filtering) [35], and R-peak detection, which is the procedure for annotating the index data points of the corresponding R-peak complex along the signal. Then, the original ECG signal is detrended to yield a better approximation for a specified length of a segment for signal analysis. The nonlinear trend in the signal is also removed by fitting a low-order polynomial to the signal and subtracting it. The polynomial is set to order 6 [36]. Subsequently, a Butterworth bandpass filter in the range of 5Hz and 15Hz is applied to eliminate the baseline wander. Baseline wanders are low-frequency noise that occurs during data acquisition in signal processing. This is usually due to perspiration that affects the electrode impedance, respiration, and body movements such as finger movements on the electrode. The detrended signal and filtering process of the original signal can be found as shown in Figure 5. The result of the filtering process is then normalized in the range of 0 and 1 by subtracting from the mean value to balance the contribution for the training phase using (13), (14) \tilde{x} , where \tilde{x} denotes the raw ECG signal and the resulting signal, respectively.

$$\tilde{x} = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (13)$$

$$\tilde{x} = \tilde{x} - \text{mean}(\tilde{x}) \quad (14)$$

Since the R-peak is the most prominent peak used to identify a heartbeat within a signal, R-peak detection is still necessary using the Pan-Tompkins algorithms [28,37] for annotating the respective peak points throughout the signal. Thus, every heartbeat in the original signal can be extracted. After annotation of the indices of the R-peak, a suitable number of samples before and after a given R-peak point are then sliced to segment the heartbeat of a signal, which is a vector form. For our fiducial approaches, we consider 125 samples before and after the R peak point to form a heartbeat for the dataset with the QRS complex, and 150 samples for the ECG-ID dataset. The other datasets are also conducted using the same approach depending on their sampling rates, and a sample of a vector that interprets the heartbeat. For each signal for every dataset, approximately 45 to 50 heartbeat segments were withdrawn using 251 samples, whereas 51 heartbeats were extracted using 301 samples from the ECG-ID dataset and other datasets for heartbeat segmentation.

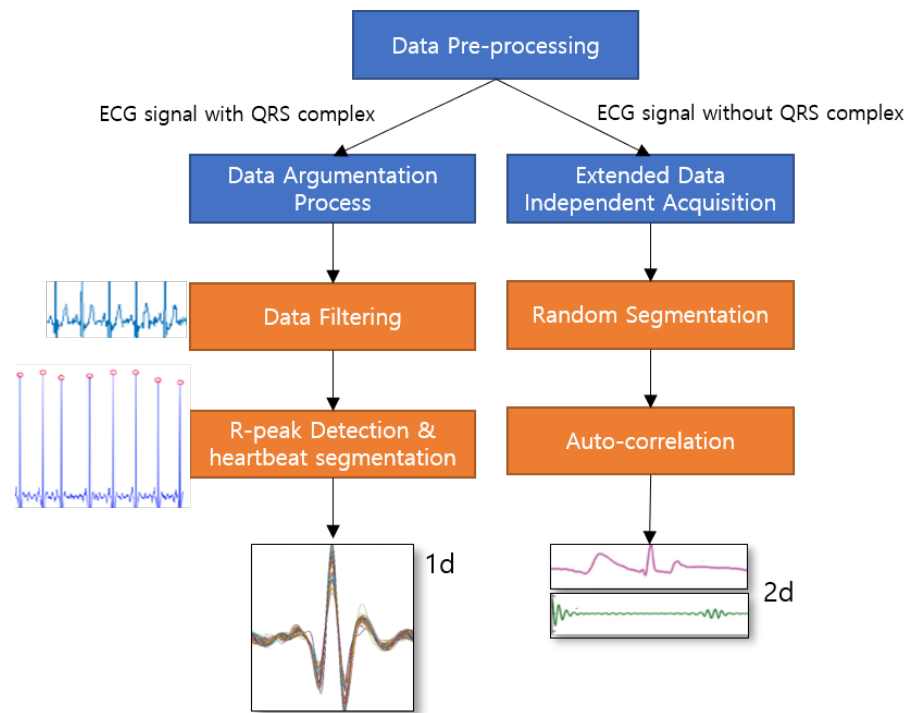


Figure 4. A framework of data preprocessing phase to carry out the proposed models according to respective datasets.

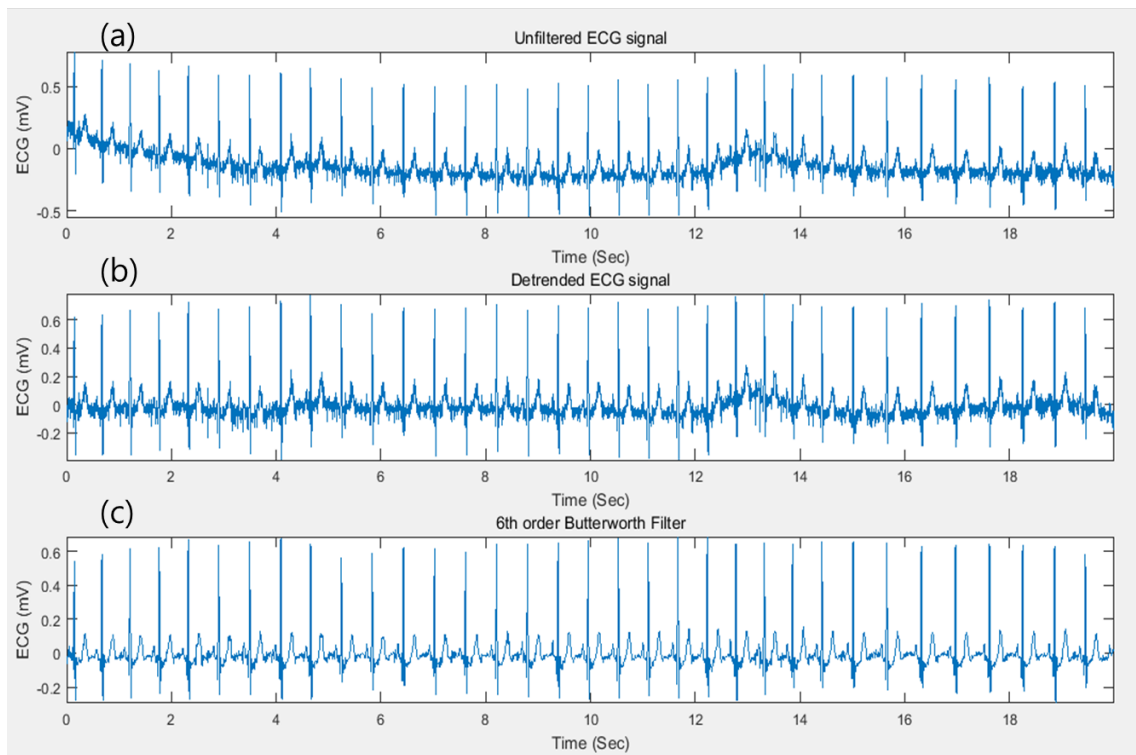


Figure 5. (a) A raw ECG signal with non-linear trend. (b) Detrended ECG signal. (c) Filtered ECG signal by applying 6th order Butterworth filter.

3.2. Extended Data Independent Acquisition

Unlike the data acquisition technique used in the previous chapter, we utilize random segmentation without applying QRS peak detection, which is available for any type of ECG signal or the signal dataset with no QRS complex. The original signal is blindly

segmented into segments with an equal length, which is a 2-s window (720 samples) to include at least one heartbeat, since the normal range for heart rate in a signal is from 40 to 280 beats per minute [25]. For each recording, 500 random windows were chosen, half of which were used to train and the other half for testing. The auto-correlation operation is introduced to remove the phase difference due to blind segmentation. The auto-correlation operation is applied to the segmented windows to remove the phase difference that occurs due to random segmentation and thus provides a shift-invariant multi-resolution data representation used in [5], and it is defined as:

$$Z_j^i[t] = \sum_{m=0}^{T-t-1} Y_j^i[m]Y_j^i[m+t], \forall t \in [0, T-1], \forall i \in [1, W]. \tag{15}$$

where $Z_j^i[t]$ is the t th sample in the j th wavelet component of the i th ECG window after auto-correlation, $Y_j^i[m+t]$ corresponds to $Y_j^i[t]$ with a time lag of m , m is chosen from 0 to $T-t-1$. T and W correspond to the number of samples in an ECG window, which is 720, and the number of ECG window which is 500, respectively.

The autocorrelation function calculates the correlation of a series with its delayed copy, i.e., the similarity between series as a function of the time lag between them [38]. Therefore, it can effectively discover repeating patterns in the quasi-periodic ECG signals even with different numbers and occurrence times of the heartbeats. After removing the phase difference, the multi-resolution data can now be fed to the networks for automatic feature learning and user identification. Figure 6 shows similar outputs when auto-correlation is applied to two different wavelet domain signal segments.

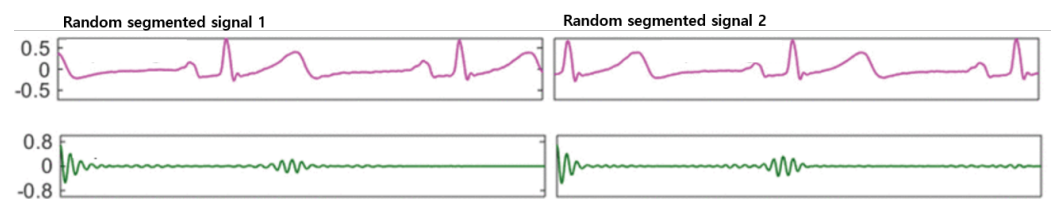


Figure 6. Multi-resolution representation of two randomly segmented windows of an ECG signal.

3.3. Models Overview

In this section, we will review several proposed techniques based on deep neural networks for downstreaming approaches. The first model is designed based on a 1-D convolutional neural network (1D-CNN), which facilitates the learning of hierarchical distinct features to present a new version for representation of a high-level abstraction. Then, such abstracted data are fed into a classification layer such as a fully connected layer for further authentication processes. The rest of the other proposed methods based on RNN with modified cell units, LSTM and GRU, are also collectively proposed and investigated, see in Figure 7. We also consider changes in their hidden states during the training procedure based on deployment in a bidirectional manner. For conventional RNN models, the hidden state of a given time step is calculated using a linear combination of the previous hidden state and the current input. Although GRU and LSTM networks share a similar network structure, the update gate of the hidden state is more complex in both approaches. Figure 8 illustrates the proposed RNN-based models for the different cell units used during the experiments in this study.

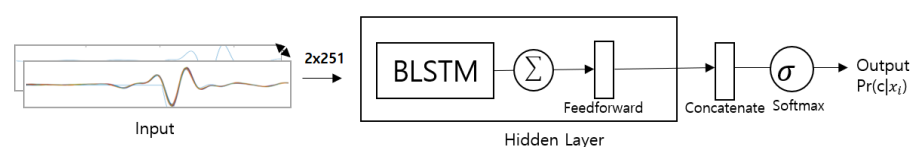


Figure 7. Overview of multi-resolution bi-directional LSTM based system architecture for ECG signal with 250 samples in a segment.

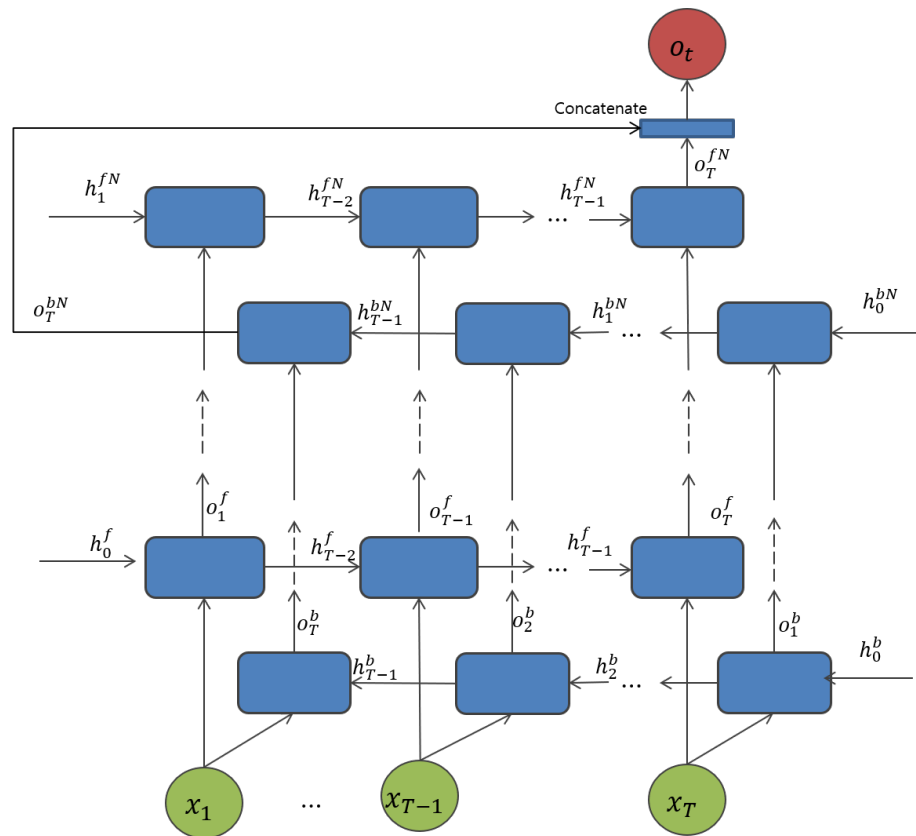


Figure 8. Proposed bi-directional RNN network for both LSTM and GRU cell unit.

3.3.1. Proposed 1-D CNN Model

Convolutional neural networks (CNNs) are neural networks built to primarily classify images, cluster images based on similarity, and perform object recognition. It was developed in the 1980s. The CNN is designed for robust training in terms of the stochastic gradient descent algorithm for each layer. Moreover, CNNs have been commonly used for feature learning and classification problems. In this report, a deep 1-D CNN is designed to perform ECG classification using a fiducial approach. The optimal parameters of the model structure are selected according to the previous works and after various lengths of trails. Moreover, all the parameters of the networks are also conducted with various settings by trial-and-error approach, and chose the optimal setting for each network which yields the most with better performance results. The weights in the models were initialized randomly at the start of the training process, and progressively updated throughout the process. The detailed network architecture of the proposed CNN model is implemented, and the respective parameters of the network used in this study are listed in Table 1.

Table 1. Values of parameters for proposed 1-D CNN model.

Layers	Kernel Size	Stride	Padding	Input Size	Output Size
1	5	2	2	750	375
2	2	2	0	375	187
3	5	2	2	187	94
4	2	2	0	94	47

There are four hidden layers in the model that are used for feature learning, followed by a fully-connected layer with 40 neurons. Then, a decision-making classification layer with a sigmoid function is applied to produce the appropriate categorical distribution for each class. The intuition of the proposed CNN model is that it allows a function that differentiates the patterns and distinct characteristics of all classes based on their respective input signal. Generally, the ground truth result is indicated in a one-hot distribution vector,

whereas the input is a discrete sequential sample, where the data point is a vectorized representation of individual samples at time t . The signals are segmented into specified windows of definite length followed by the procedure presented in the first section of this report. Each window captures at least one or more heartbeat waveforms in the original signal. The parameter values, such as filter size, stride, and padding values, are set according to Table 1 throughout the network layers from the first layer to the last layer. Convolutional operations with nonlinear activation functions are applied between each layer. In the first and second layers, 30 filters are implemented. Finally, the softmax function is used in the last layer to produce the distribution of the corresponding class for decision making in the form of a vector, in the range of 0 to 1. The cross-entropy loss function for the network's targets can be calculated as

$$E = - \sum_{i=1} (\hat{y}_i \log(y_i) + (1 - \hat{y}_i) \log(1 - y_i)) \quad (16)$$

where \hat{y}_i is the ground truth target vector and y_i is the output vector of our model for class i . To obtain the output in categorical distributions across all the subjects, outputs $y_{1,\dots,i}$ are calculated by applying a sigmoid function to the weight sums of the activation function of the previous layer.

3.3.2. Proposed Bidirectional RNN Architectures

To investigate the use of RNN methods for ECG classification, we proposed the models based on different types of RNN techniques. The input training data can be set by $S = (X_n, O_n), n = 1, \dots, N$, where the sample suggests the m numbers of samples in a signal with 251-sampled segments following the procedure used in the preprocessing phase, whereas 301-sampled segmented windows are for the ECG-ID dataset, where i denotes the number of O_i^n subject, which denotes the corresponding ground truth for each subject of the n th input. The ground truth values are determined as 1 for a given subject signal, and 0 for the other subjects, respectively. For a given sequence input, a classifier is trained to learn the probabilities of N classes.

The first proposed model for the non-fiducial approach is based on a bidirectional RNN with an LSTM cell unit in the hidden state layer, and is called BLSTM, as shown in Figure 3a. It is associated with the cell unit. The segmented signal inputs x_1, \dots, x_T from the pre-processing section, are fed into the network for each time step $t (t = 1, \dots, T)$ for each LSTM cell. Each cell unit in a bi-directional manner consists of a parallel of LSTM tracks, known as forward and backward sequences, to capture the context from the past and future. During the final time step, the two parallel tracks of the LSTM cell unit are concatenated into a single vector. In the first hidden layer, the forward cell states h_0^f , and the backward cell state h_0^b are initialized with zero for all layers N . The input x_t at time t , and previous cell states h_{t-1} to produce the output of the corresponding layer p_t^n at time t and at the n th layer for both backward or forward tracks given its parameter θ^n , can be defined as

$$o_t^n, h_t^n = LSTM^n(h_{t-1}^n, x_t; \theta^n) \quad (17)$$

$$o_t^n, h_t^n = GRU^n(h_{t-1}^n, x_t; \theta^n) \quad (18)$$

where θ^n denotes the parameters (b, U, W) of the respective cell unit for layer n .

For the next proposed model, see Figure 3b, the only difference between BLSTM and the GRU gate units is the cell unit at the hidden layers. In addition, to address one of the most important challenges in deep neural networks, overfitting, the dropout layer is applied to each cell for all RNN-based methods. Sharing the similar outputs as in BLSTM at the last layer, the outputs from both forward and backward track, the late fusion for bidirectional networks is concatenated into a single vector. Then, the output is followed by a softmax activation function to achieve the N -dimensional output in the last layer. The overall model architecture is shown in Figure 7. As implemented based on a bi-directional manner, the forward track trains the input from left to right, while the

backward track traces the input from right to left in both BLSTM and BGRU, and can be defined as follows:

$$o_t^f, h_t^f, c_t^f = LSTM^f(c_{t-1}^f, h_{t-1}^f, x_t; W^f) \quad (19)$$

$$o_t^b, h_t^b, c_t^b = LSTM^b(c_{t-1}^b, h_{t-1}^b, x_t; W^b) \quad (20)$$

$$o_t^f, h_t^f, c_t^f = GRU^f(c_{t-1}^f, h_{t-1}^f, x_t; W^f) \quad (21)$$

$$o_t^b, h_t^b, c_t^b = GRU^b(c_{t-1}^b, h_{t-1}^b, x_t; W^b) \quad (22)$$

4. Experimental Results and Discussion

4.1. Network Training

To achieve a higher acceleration of the training process, which is usually a bottleneck while operating deep networks with many layers, our proposed models are developed in the Tensorflow deep learning library, which can be executed on a graphics processing unit (GPU). It is commonly at least 5 to 10 times faster than the central processing unit (CPU), and can also predominantly increase the training process. All our experiments were executed on a GeForce GTX 1080 GPU. During the training process, the 1D-CNN based model learns hierarchical features by performing convolution and pooling operations in accordance with the parameters provided in Table 1. The stochastic gradient descent (SGD) learning method is applied to increase the acceleration of the training process. This allows for passing a batch of training input data to the neural network each time. The batch size is selected as 150 for all proposed methods, including RNN-based networks, to achieve a compromise based on two considerations. Specifically, a large size results in a small convergence time by reducing the variance of stochastic gradient updates, and a small size to strengthen SGD to leap out the shallow minima during the error loss function. However, the network can learn the hidden patterns of the input signal and reach its convergence at 14 epochs. The epoch size is set as 50 to offset the under-fitting and over-fitting considerations.

For RNN-based models, the batch size is selected as 150 because it yields better performance compared to the other schemes, as previously indicated. The optimization method was applied using the Adam optimizer and the learning rate was set to 0.001. Moreover, the loss functions were determined based on the categorical cross-entropy method used, where δ_l indicates the ground truth vector, and o_l denotes the output vector of the model for the l class. In our experiments on RNN-based methods, the optimal window length of the segmented signal is chosen with regard to previous works, and after the various lengths of the attempts. The parameters of the proposed models were also examined using various trials of the settings, and the optimal setting was selected to yield better performance. The weight parameters in the proposed models were randomly initialized at the beginning of the training process and incrementally updated throughout the entire procedure. A dropout value of 0.2 for outputs in the first layer of the networks and the last layer inputs was used to avoid the over-fitting problem that is typically encountered in the learning of deep neural networks. The top block indicates the training cross-entropy loss as it reaches its convergence at 70 epochs when the percentage of subjects used for training is 50%.

$$E = - \sum_{i=1} (\delta_l \log(o_l) + (1 - \delta_l) \log(1 - o_l)) \quad (23)$$

4.2. System Evaluation

For our experiment, the ECG-ID (ECG-ID) and MIT-BIH ECG, STAFF III database (STAFF-III), and Long Term AF Database (LT-AF) data were collected from PhysioNet. The ECG-ID dataset includes 310 ECG recordings digitized at 500 Hz, obtained from 90 subjects (10,000 samples), while the MIT-BIH ECG dataset contains 168 short recordings for 47 subjects, set to pose a variety of challenges for ECG compressors, especially

for compression methods. The STAFF III database was acquired during 1995–1996 and contains standard 12-lead ECG recordings from 108 patients. LT-AF dataset consists of ECG recordings of 84 subjects with paroxysmal or sustained atrial fibrillation (AF), and digitized at 128 Hz with durations which vary from 24 to 25 hours. Moreover, two more datasets, namely AFDB with 20 subjects and AHA datasets with two subjects (without QRS complex), were examined for our extended data-independent acquisition-based approach. These datasets may be acquired by different lead configurations. Moreover, the datasets were collected from healthy or quasi-healthy participants, and also include severe heart diseases such as ST depression or elevation, arrhythmia, atrial fibrillation and malignant ventricular ectopy. However, the datasets were not acquired with the same sampling rate, all the ECG recordings were resampled to 360 Hz to fairly illustrate the performance. To train the RNN-based networks, the training dataset was divided into batches of several heartbeats for each dataset. The weights for each batch were updated upon completion. The input data was forward and backward propagated throughout the network, and the error cost was calculated by back-propagating the unfolded network in time. We adopted a method called backpropagation through time (BPTT) with the Adam optimization method employed in our experiment, with the learning rate set to 0.001. The batch size was set as 150, which yielded higher performance for all the methods, namely, traditional RNN, RNN with LSTM, and RNN with GRU. The epoch size was chosen as 150 to balance network over-fitting issues. Furthermore, a dropout [39] of 0.4 outputs of the first layer, and the last layer inputs were utilized to overcome the overfitting problem. For the evaluation, total of 5066 segmented windows from all dataset from all dataset were separated for training and test sets. To investigate the accuracy of the classification rate, the proposed models were evaluated based on classification accuracy, which can be determined using the confusion matrix. It is one of the most common intuitive metrics used for evaluating the performance and accuracy of machine learning models commonly used for classification problems, as shown in Table 2.

Table 2. Confusion matrix for evaluation classification accuracy.

		Actual	
		Positives (1)	Negatives (0)
Predicted	Positives (1)	TP	FP
	Negatives (0)	FN	TN

The associated terms are related to a given confusion matrix can be specified as true positives (TP), True negatives (TN), False positives (FP), and True negatives (TN). When the output correctly classifies the data point as the ground truth, it is considered to be a TP. True negatives (TN) are the cases when the output class of the given data point is predicted correctly as negative for the given class. False positives (FP) are the cases when the model incorrectly predicted the corresponding class as positive. False negatives (FN) are the cases when the ground truth should be positive, but the model’s prediction is negative. As a result of the accuracy in classification tasks if the correct predictions made by the model over all predictions, the correct predictions are known as true positive (TP) and true negative (TN) are divided by all predictions made by the model, calculated in (24).

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (24)$$

According to Table 3, the reported overall classification accuracy outperforms all the previous studies for all datasets. Before we investigated the performance of our proposed method, we studied it in comparison to conventional RNN-based methods, namely, traditional RNN, RNN with LSTM gates, and RNN with GRU gates over four datasets with QRS complex. The optimal input sequence length or the number of successive heartbeats for the experiments are selected according to the previous experiments and after various lengths of trails by conducting with widely used lengths; single heartbeat, three

heartbeats, six heartbeats and nine heartbeats. Since there is a lower accuracy rate in single heartbeat compared to the rest of the groups, and there is no significant higher accuracy rate in six-heartbeat group compared to three-heartbeat group, we selected the number of successive heartbeats to three and nine for further experiments.

Table 3. Performance of classification accuracy for selected input sequence length over conventional and proposed models.

Type of Model	Input Sequence Length (Number of Heartbeats)	Accuracy
Proposed 1D-CNN	3	0.925
	9	0.911
RNN + LSTM	3	0.965
	9	0.971
RNN + GRU	3	0.952
	9	0.978
Proposed BLSTM	3	0.982
	9	0.993
Proposed BGRU	3	0.921
	9	0.983

The entropy loss plot is shown in Figure 9, and the proposed multi-resolution bi-directional LSTM outperforms the others as the training process reaches 150 epochs over the ECG-ID dataset. Thus, our proposed bi-directional LSTM-based model was chosen for further experiments that examined our new data-independent acquisition-based method compared to other recent studies.

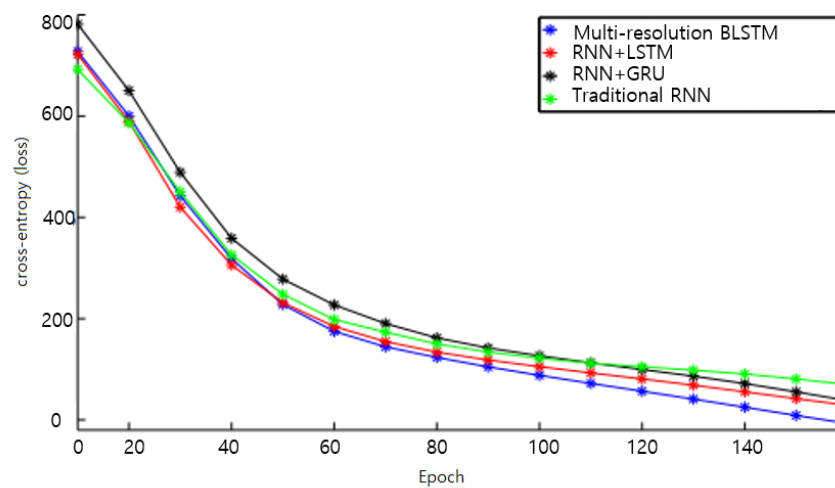


Figure 9. Comparison of cross-entropy loss per epoch over ECG-ID dataset over conventional and proposed models.

To evaluate the performance results for classification, we examined the networks based on four statistical evaluation metrics: accuracy (*Acc*), sensitivity (*Sen*), specificity (*Spc*), and positive predictivity (*Ppr*).

$$Acc = \frac{TP + TN}{TP + FP + FN + TN} \quad (25)$$

$$Sen = \frac{TP}{TP + FN} \quad (26)$$

$$Spe = \frac{TN}{TN + FP} \quad (27)$$

$$Ppr = \frac{TP}{TP + FP} \quad (28)$$

The terms TP , TN , FP , and FN in the preceding equations denote true positive, true negative, false positive, and false negative, respectively. $F1$ and Fowlkes-Mallows index (FM) scores are also computed using Sen and Ppr as follows:

$$F1 = \frac{2}{\frac{1}{Sen} + \frac{1}{Ppr}} \quad (29)$$

$$FM = \sqrt{Sen \times Ppr} \quad (30)$$

We compared the proposed method with previous RNN-based networks for six ECG datasets according to the preceding evaluation metrics. Our proposed multi-resolution bidirectional LSTM outperformed the others in terms of $F1$ and FM scores for the ECG-ID dataset. The accuracy of the proposed algorithm for both datasets nearly matches that of Mostayed et al. [24], which used a bidirectional LSTM network for a 12-lead ECG signal. It should be noted that the $F1$ score is a more significant metric than the accuracy score. Q. Zhang et al. [5] used a multi-resolution parallel network based on CNN, deploying multiple versions of wavelets to improve the context representation of the signal for generalization purposes. For the ECG-ID dataset, the accuracy for all techniques was quite similar, and the proposed method outperformed the other methods for the $F1$ score, which achieved 98.84%. For the MIT-BIH ECG dataset, Fan Liu et al. [40] achieved a higher accuracy compared to other techniques [24,40,41]; however, our proposed method significantly outperformed theirs. Moreover, they still need to first identify the heartbeat, which is intensive in terms of the algorithm engineering process, compared to our random signal segmentation. In the STAFF-III dataset, the proposed method scored a high accuracy of 97%, and achieved higher $F1$ and FM scores. Regarding the LT-AF dataset, despite obtaining a similar accuracy score to A. Mostayed et al. [24] of 99%, our proposed method significantly outperformed the others in terms of $F1$ and FM scores of 99.5%.

However, for two datasets with signals without a QRS complex, the rest of the compared techniques based on the R-peak detection method for heartbeat segmentation significantly decreased in terms of classification accuracy. The highest $F1$ score did not reach 90.5% among all the compared techniques, whereas our data-independent acquisition approach achieved scores of 97.94% and 97.3% for the AFDB and AHA datasets, respectively. Nevertheless, the proposed method has more advantages and a sufficiently high classification capability compared to the recent studies in the overall comparison of different signal versions.

Figures 10–12 show the training process for all six trials. Interestingly, we found that our proposed method had the best convergence speed and least training entropy loss after 120 epochs for the ECG-ID dataset, and it reached its convergence after 90 epochs for the MIT-BIG ECG dataset, while the epoch rates decreased to 60 to 90 for both the STAFF-III and LT-AF datasets, which are considered as presenting better signal quality with less noise and shift invariants. However, in both datasets without the QRS complex, AFDB, and AHA, none of the other compared methods using QRS peak detection were unable to achieve a high convergence speed even after 170 epochs. This shows that data-dependent methods can only perform well for specified data types and amendments in data variations. Figure 13 showed Receiver Operating Characteristics (ROC) analysis of the proposed bi-directional LSTM network on different ECG datasets provided Sensitivity and Specificity for different datasets are respective to Table 4, and it suggests that the proposed model yields better classification on MIT-BIH and ECG-ID with QRS complex compared to the rest.

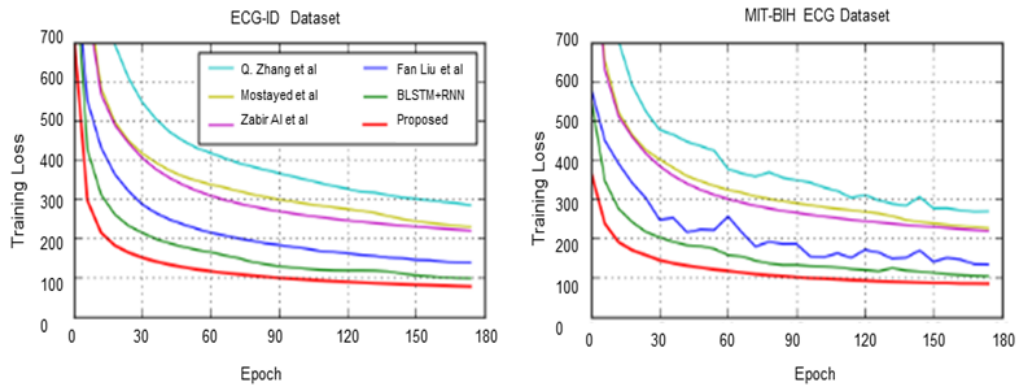


Figure 10. Comparison of cross-entropy loss per epoch over ECG-ID dataset and MIT-BIH ECG dataset.

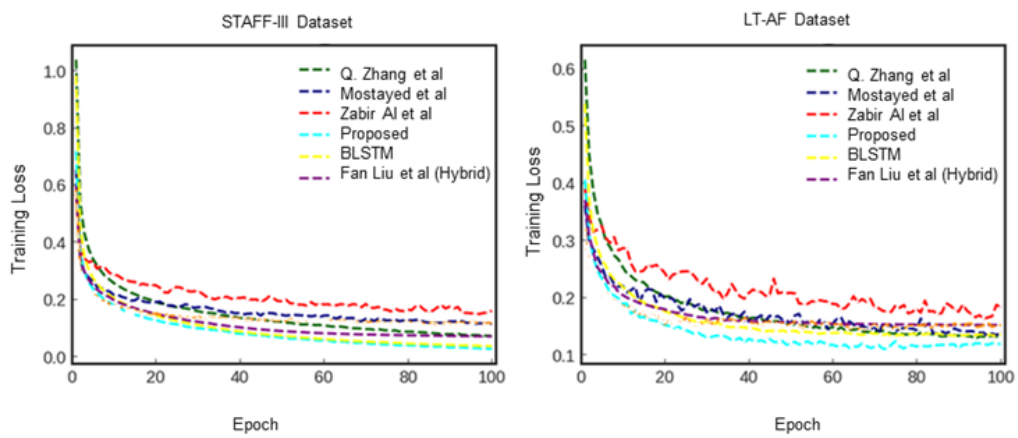


Figure 11. Comparison of cross-entropy loss per epoch over STAFF-III dataset and LT-AF dataset.

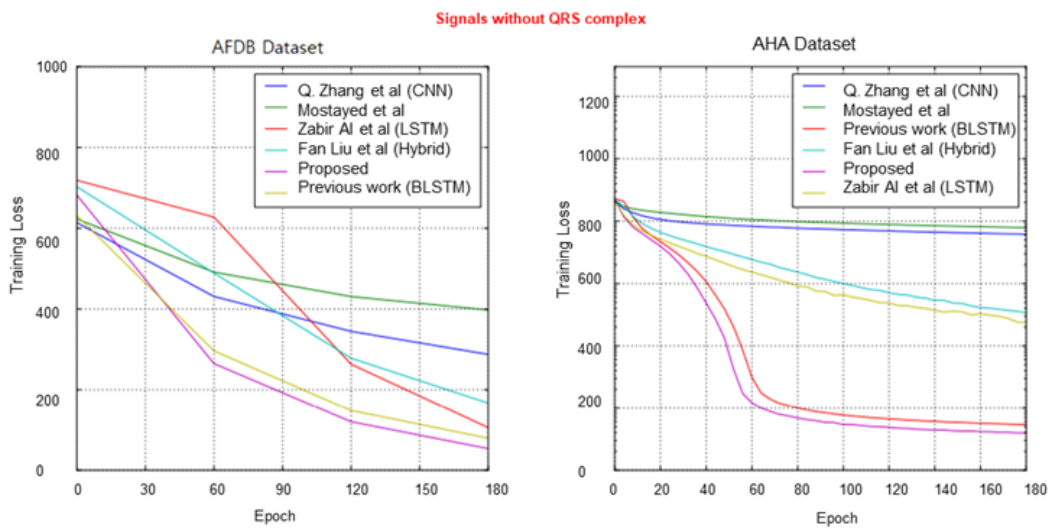
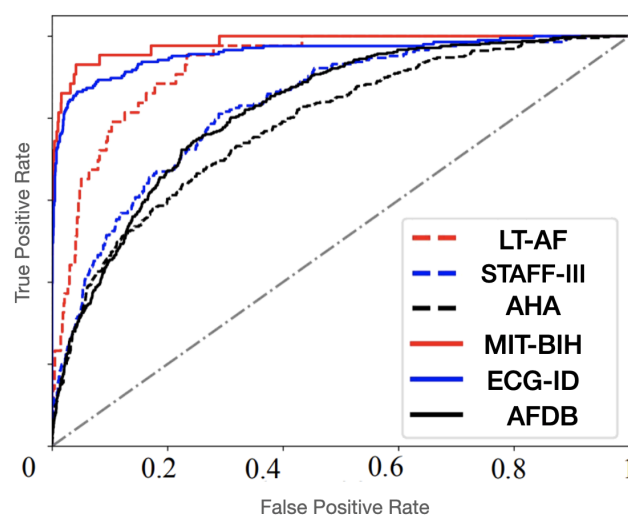


Figure 12. Comparison of cross-entropy loss per epoch over AFDB dataset and AHA dataset.

Table 4. Comparison of the proposed method with recent state-of-the-art works in terms of classification accuracy of 20 classes over four different datasets (ECG-ID, MIT-BIH, STAFF-III, LT-AF, AFDB, and AHA).

Dataset	Method	Acc	Sen	Spe	Ppr	F1	FM
ECG-ID	Zhang et al. [5]	98.3	75.2	98.3	99.8	85.8	85.8
	Mostayed et al. [24]	98.4	93	97.5	98.2	95.5	95.5
	Zabir Al et al. [41]	90.3	94.2	95.6	93.1	93.6	93.6
	Fan Liu et al. [40]	98.3	95.7	98.2	99.2	97.4	97.4
	Proposed	99.3	98.3	99.2	99.4	98.84	98.84
MIT-BIH ECG	Zhang et al. [5]	98.6	95.2	97.3	89.5	92.2	92.2
	Mostayed et al. [24]	99.4	95.8	99.7	97.8	96.8	96.8
	Zabir Al et al. [41]	80.1	82.8	89.1	84.4	83.59	83.59
	Fan Liu et al. [40]	80.1	82.8	89.1	84.4	83.59	83.59
	Proposed	99.5	99.2	98.8	99.2	99.2	99.2
STAFF-III	Zhang et al. [5]	98.1	86.6	99.3	96.2	91.2	91.2
	Mostayed et al. [24]	98.7	91.3	97.4	97.8	94.4	94.4
	Zabir Al et al. [41]	89.4	88.7	92.3	89.6	89.14	89.14
	Fan Liu et al. [40]	98.6	94.6	99.2	98.5	96.5	96
	Proposed	99.3	95.5	97.9	99.2	97.31	97.07
LT-AF	Zhang et al. [5]	97.6	95.8	95.3	96.4	96	96
	Mostayed et al. [24]	99.1	99.4	98.7	98.5	98.6	98.6
	Zabir Al et al. [41]	89.4	88.4	90.2	93.2	90.7	90.76
	Fan Liu et al. [40]	99.4	99.2	98.4	97.6	98.39	98.39
	Proposed	99.2	99.6	98.2	99.5	99.5	99.5
AFDB	Zhang et al. [5]	89.6	91.3	90.5	88.7	90	90
	Mostayed et al. [24]	83.1	88.4	89.3	88.3	88.34	88.34
	Zabir Al et al. [41]	79.1	81.2	88.4	86.5	83.8	83.8
	Fan Liu et al. [40]	90.2	89.8	92.5	90.3	90.4	90.4
	Proposed	98.5	97.3	99.1	98.6	97.94	97.94
AHA	Zhang et al. [5]	84.3	81.5	83.6	88.4	84.87	84.87
	Mostayed et al. [24]	85.4	86.4	88.3	83.4	84.8	84.8
	Zabir Al et al. [41]	76.5	81.2	83.2	88.2	84.62	84.62
	Fan Liu et al. [40]	87.5	88.3	89.2	83.6	85.91	85.91
	Proposed	97.3	96.4	97.2	98.4	97.39	97.39

**Figure 13.** Receiver Operating Characteristics (ROC) classification of the proposed bi-directional LSTM network for six different ECG datasets.

This is consistent with the conceptual experiment that establishes that poor local minima are hardly an issue in deep neural networks with many layers consisting of a large number of parameters. Instead, the landscape of the object function is packed with a variation of valleys, which seems to typically have local minima with similar values. Therefore, the randomness in the Adam optimization-based method parameter tuning process often results in only small fluctuations in the convergence curve in the training process.

4.3. Discussion

We still need to put more effects in several aspects compared to the current well-known biometric techniques such as fingerprint identification [42]. For instance, although more and more ECG datasets are available now benefitting from great advancement of data acquisition systems, the amount of data is still not comparable to the fingerprint data. Generally, hundreds or thousands of records of fingerprint can be obtained, which is much higher than the dataset sizes used in this study. Moreover, there are also more diversity induced by different heart health conditions, which brings more challenges to the generalization ability of the identification algorithms [43,44]. Our proposed method includes random segmentation for data enriching purpose as non-fiducial approach, which effectively avoids complicated and data-dependent signal preprocessing efforts such as QRS peaks detection and segmentation effort. It also removes phase difference among random-chosen signal segments by the auto-correlation approach. However, further consideration for more data representation methods and deep neural network techniques should be considered, to explore the interesting connection between the data representation network topology and the feature learning ability. Our experimental results have also suggested that the cardiac organisms react physiologically to stimulations such as danger and threat, and corresponding ECG signal also changes due to the cardiac defense mechanism (CDR). Therefore, identifying these external stimulation-related ECG patterns in future work is also expected to further improve the user identification performance.

5. Conclusions and Future Work

We proposed a multi-resolution bidirectional LSTM network with a random segmentation technique using an auto-correlation method to increase the dimension of the input ECG data for biometric identification. By applying random segmentation with an auto-correlation approach for independent data acquisition, the time-frequency representation of the original signal was improved, and the classification accuracy for data variations increased. The performance of the learning procedure was improved compared to other RNN-based methods and hybrid method based approaches according to the experimental results. The experimental outcomes showed that the proposed algorithm outperformed most RNN based networks by adapting the bidirectional learning method, and considerably improving the classification performance for more contextualized distinct features. We will further consider more data representation techniques and deep learning methods for better feature learning capability, and demonstrating the difference between traditional RNN networks and proposed BLSTM method. The proposed method of the BLSTM-RNN model can also be considered for generalization to other periodic waveforms for biometric signal-based user authentication applications.

Author Contributions: Conceptualization, H.M.L.; methodology, H.M.L. and S.B.P.; formal analysis, H.M.L.; writing—original draft preparation, H.M.L.; review and editing, S.B.P. and P.K. All authors have read and agreed to the published version of the manuscript.

Funding: Please refer to Acknowledgements.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(NRF-2017R1A6A1A03015496).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Stavroulakis, P.; Stamp, M. *Handbook of Information and Communication Security*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2010.
2. Sufi, F.; Khalil, I.; Hu, J. ECG-based authentication. In *Information and Communication Security*; Springer: Berlin/Heidelberg, Germany, 2010; pp. 309–331.
3. Acharya, U.R.; Oh, S.L.; Hagiwara, Y.; Tan, J.H.; Adam, M.; Gertych, A.; Tan, R.S. A Deep convolutional neural network model to classify heartbeats. *Comput. Biol. Med.* **2017**, *89*, 389–396. [[CrossRef](#)]
4. Yildirim, Ö. A novel wavelet sequence based on a deep bidirectional LSTM network model for ECG signal classification. *Comput. Biol. Med.* **2018**, *96*, 189–202. [[CrossRef](#)]
5. Zhang, Q.; Zhou, D.; Zeng, X. HeartID: A multiresolution convolutional neural network for ECG-based biometric human identification in smart health applications. *IEEE Access* **2017**, *5*, 11805–11816. [[CrossRef](#)]
6. Pan, J.; Tompkins, W.J. Real-time QRS detection algorithm. *IEEE Trans. Biomed. Eng.* **1985**, *32*, 230–236. [[CrossRef](#)]
7. Raman, P.; Ghosh, S.M. Classification of heart diseases based on ECG analysis using FCM and SVM methods. *Int. J. Eng. Sci.* **2016**, *6*, 673–6744.
8. Sahoo, S.; Kanungo, B.; Behera, S.; Sabut, S. Multiresolution wavelet transform-based feature extraction and ECG classification to detect cardiac abnormalities. *Measurement* **2017**, *108*, 55–66. [[CrossRef](#)]
9. Thomas, M.; Das, M.K.; Ari, S. Automatic ECG arrhythmia classification using dual tree complex wavelet-based features. *AEU-Int. J. Electron. Commun.* **2015**, *69*, 715–721. [[CrossRef](#)]
10. Martis, R.J.; Acharya, U.R.; Min, L.C. ECG beat classification using PCA, LDA, ICA, and discrete wavelet transform. *Biomed. Signal Process. Control.* **2013**, *8*, 437–448. [[CrossRef](#)]
11. LeCun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning is applied to document recognition. *Proc. IEEE* **1998**, *86*, 2278–2324. [[CrossRef](#)]
12. Srivastava, R.K.; Greff, K.; Schmidhuber, J. Highway networks. *arXiv* **2015**, arXiv:1505.00387.
13. Bengio, Y. Learning deep architectures for AI. *Found. Trends Mach. Learn.* **2009**, *2*, 1–127. [[CrossRef](#)]
14. Yildirim, Ö.; Uçar, A.; Baloglu, U.B. Recognition of Real-world Texture Images under Challenging Conditions with Deep Learning. In Proceedings of the ASYU, Alanya, Turkey, 5–7 October 2017; Volume 52.
15. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* **2012**, *60*, 1097–1105. [[CrossRef](#)]
16. Uçar, A.; Demir, Y.; Güzeliş, C. Object recognition and detection with deep learning for autonomous driving applications. *Simulation* **2017**, *9*, 759–769. [[CrossRef](#)]
17. Min, S.; Lee, B.; Yoon, S. Deep learning in bioinformatics. *Briefings Bioinf.* **2017**, *18*, 851–869. [[CrossRef](#)]
18. Tan, J.H.; Hagiwara, Y.; Pang, W.; Lim, I.; Oh, S.L.; Adam, M.; San Tan, R.; Chen, M.; Acharya, U.R. Application of stacked convolutional and long short-term memory networks for accurate identification of CAD ECG signals. *Comput. Biol. Med.* **2018**, *94*, 19–26. [[CrossRef](#)]
19. Kiranyaz, S.; Ince, T.; Gabbouj, M. Real-time patient-specific ECG classification by 1-D convolutional neural networks. *IEEE Trans. Biomed. Eng.* **2016**, *63*, 664–675. [[CrossRef](#)]
20. Zhang, X.; Zhang, Y.; Zhang, L.; Wang, H.; Tang, J. Ballistocardiogram-based Person Identification and Authentication Using Recurrent Neural Networks. In Proceedings of the 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics, Beijing, China, 13–15 October 2018; pp. 1–5.
21. Warrick, P.; Homsy, M.N. Cardiac Arrhythmia Detection from ECG Combining Convolutional and Long Short-term Memory Networks. In Proceedings of the 2017 Computing in Cardiology (CinC), Rennes, France, 24–27 September 2017.
22. Zihlmann, M.; Perekrestenko, D.; Tschannen, M. Convolutional recurrent neural networks for electrocardiogram classification. In Proceedings of the 2017 Computing in Cardiology (CinC), Rennes, France, 24–27 September 2018.
23. Coşkun, M.; Uçar, A.; Yildirim, Ö.; Demir, Y. Face recognition based on a CNN. In Proceedings of the 2017 International Conference on Modern Electrical and Energy Systems (MEES), Kremenchuk, Ukraine, 15–17 November 2017; pp. 376–379.
24. Mostayed, A.; Luo, J.; Shu, X.; Wee, W. Classification of 12-Lead ECG signals with bi-directional LSTM network. *arXiv* **2018**, arXiv:1811.02090
25. Tanaka, H.; Monahan, K.D.; Seals, D.R. Age-predicted maximal heart rate was revisited. *J. Am. Coll. Cardiol.* **2001**, *37*, 153–156. [[CrossRef](#)]
26. Cho, K.; Van Merriënboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y. Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv* **2014**, arXiv:1406.1078.
27. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [[CrossRef](#)] [[PubMed](#)]
28. Lugovaya, T.S. Biometric Human Identification Based on Electrocardiogram. Master’s Thesis, Faculty of Computing Technologies and Informatics, Electrotechnical University, Saint Petersburg, Russia, 2005.

29. Moody, G.B.; Mark, R.G. Impact of the MIT-BIH Arrhythmia Database. *IEEE Eng. Med. Biol.* **2001**, *2*, 45–50. [[CrossRef](#)] [[PubMed](#)]
30. Martinez, J.P.; Pahlm, O.; Ringborn, M.; Warren, S.; Laguna, P.; Sornmo, L. The STAFF III Database: ECGs Recorded During Acutely Induced Myocardial Ischemia. *Comput. Cardiol.* **2017**, *44*, 1–4.
31. Petrutiu, S.; Sahakian, A.V.; Swiryn, S. Abrupt changes in fibrillatory wave characteristics at the termination of paroxysmal atrial fibrillation in humans. *Europace* **2007**, *9*, 466–470. [[CrossRef](#)] [[PubMed](#)]
32. Moody, G.B. Spontaneous Termination of Atrial Fibrillation: A Challenge from PhysioNet and Computers. *Comput. Cardiol.* **2004**, *31*, 101–104.
33. Goldberger, A.; Amaral, L.; Glass, L.; Hausdorff, J.; Ivanov, P.C.; Mark, R.; Stanley, H.E. PhysioBank, PhysioToolkit, PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation* **2010**, *101*, 215–220. [[CrossRef](#)]
34. Goldberger, A.L.; Amaral, L.A.; Glass, L.; Hausdorff, J.M.; Ivanov, P.C.; Mark, R.G.; Mietus, J.E.; Moody, G.B.; Peng, C.K.; Stanley, H.E. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiological signals. *Circulation* **2000**, *101*, e215–e220.
35. MathWorks. Remove Trends from Data. 2018. Available online: <https://www.mathworks.com/help/signal/ug/remove-trends-from-data.html> (accessed on 9 November 2020).
36. Al Rahhal, M.M.; Bazi, Y.; AlHichri, H.; Alajlan, N.; Melgani, F.; Yager, R.R. Deep learning approach for the active classification of ECG signals. *Inf. Sci.* **2016**, *345*, 340–354. [[CrossRef](#)]
37. Hamilton, P.S.; Tompkins, W.J. Quantitative investigation of QRS detection rules using the MIT/BIH Arrhythmia Database. *IEEE Trans. Biomed. Eng.* **1986**, *BME-33*, 1157–1165. [[CrossRef](#)]
38. Elman, J.L. Finding structure in time. *Cognit. Sci.* **1990**, *14*, 179–211. [[CrossRef](#)]
39. Srivastava, N.; Hinton, G.E.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929–1958.
40. Liu, F.; Zhou, X.; Wang, T.; Cao, J.; Wang, Z.; Wang, H.; Zhang, Y. Attention-based Hybrid LSTM-CNN Model for Arrhythmias Classification. In Proceedings of the 2019 International Joint Conference on Neural Networks, Budapest, Hungary, 14–19 July 2019; pp. 1–8.
41. Nazi, Z.A.; Biswas, A.; Rayhan, M.A.; Azad Abir, T. Classification of ECG signals by dot residual LSTM network with data augmentation for anomaly detection. In Proceedings of the 22nd International Conference on Computer and Information Technology, Dhaka, Bangladesh, 18–20 December 2019.
42. Ali Salah, A. Machine learning for biometrics. In *Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques*; IGI Global: Oxford, UK, 2010; Chapter 26, pp. 539–560.
43. Yu, S.N.; Chou, K.T. Integration of independent component analysis and neural networks for ECG beat classification. *Expert Syst. Appl.* **2008**, *34*, 2841–2846. [[CrossRef](#)]
44. Inan, O.T.; Giovangrandi, L.; Kovacs, G.T. Robust neural-network-based classification of premature ventricular contractions using wavelet transform and timing interval features. *IEEE Trans. Biomed. Eng.* **2006**, *53*, 2507–2515. [[CrossRef](#)] [[PubMed](#)]