

Improving ECG SVEB Detection using EMD with Resampling based on LSTM*

AGUNG TRI ARTA HARIYANTO¹, KUOCHEN WANG^{1,2,+}
WEI-CHUNG TSAI^{3,4} AND CHI-TSUNG CHEN⁵

¹*Department of Computer Science*

National Yang Ming Chiao Tung University

²*Center for Fundamental Science*

⁵*Office of Library and Information Services*

⁴*Faculty of Medicine, College of Medicine
Kaohsiung Medical University*

E-mail: kwang@kmu.edu.tw⁺

³*Division of Cardiology*

Department of Internal Medicine

Kaohsiung Medical University Hospital

Atrial Fibrillation (AF) is the most common arrhythmia type that affects patients today. Detecting and classifying a patient's electrocardiogram (ECG) beats, especially the supra-ventricular ectopic beats (SVEB) class, can help assess if the patient has high possibilities of AF/atrial flutter in the future. Detecting the SVEB class considered more difficult than the other classes. Related works show low classification (prediction) performance, in terms of sensitivity, F1 score, and G score, for detecting the SVEB class in a single-lead ECG. This work focuses on designing an arrhythmia beats detection method using single-lead ECG data with a patient-specific training model design, and does classification based on the AAMI standards. This work aims at achieving high classification performance in the SVEB class and still meets the real-time ECG classification requirement. The proposed method uses Empirical Mode Decomposition (EMD) with resampling (EMDR), which resamples only the first Intrinsic Mode Function (*i.e.*, IMF 1) as a main input, for the proposed EMDR-LSTM (Long Short-Term Memory) architecture. In contrast to the related works that use two separate models with one or two LSTM layers for each input, we designed a novel LSTM model architecture that only uses a single model with one LSTM layer for each input. The proposed LSTM architecture is suited for our preprocessing method, EMDR, and can enhance the SVEB classification performance. To the best of our knowledge, the proposed EMDR-LSTM is the first one that uses resamples first IMF in LSTM that classifies arrhythmia using single-lead ECG data based on the AAMI standards. Compared to representative related works, experiment results show that the proposed EMDR-LSTM achieves the highest classification performance, in terms of the following performance metrics: accuracy, sensitivity, positive predictivity, and F1 and G scores, for the SVEB class in all datasets used. In addition, although the proposed EMDR-LSTM has higher preprocessing cost and higher computational complexity in terms of MACs (multiply-accumulate operations), it has lower standard deviations of the performance metrics and lower inference time, which are important performance metrics for real time or time-critical applications, *e.g.*, ECG medical monitoring applications, compared to the representative related works.

Keywords: Atrial Fibrillation (AF), ECG classification, empirical mode decomposition with resampling (EMDR), long short-term memory (LSTM), supraventricular ectopic beats (SVEB)

Received March 19, 2021; revised September 19, 2021; accepted December 19, 2021.

Communicated by Tzung-Pei Hong.

⁺ Corresponding author.

* It was supported by the Ministry of Science and Technology (MOST), Taiwan under Grants MOST 108-2221-E-009-052 and MOST 110-2221-E-037-006 and by the Kaohsiung Medical University Hospital under Grants KMH105-5M07 and KMH106-6R10.

1. INTRODUCTION

Atrial Fibrillation (AF) is the most common arrhythmia type that affects patients today. Patients with AF are at risk of heart failure, angina, dementia, and stroke [1, 21]. Disease prevalence of AF is increasing at an alarming rate worldwide [21, 22]. Overall, in patients with AF, the crude mortality rate for all-cause death was 63.3 per 1,000 person-years [29]. The estimated number of individuals with AF globally in 2010 was 33.5 million (20.9 million men and 12.6 million women) and is predicted to affect 6-12 million people in the USA by 2050 and 17.9 million in Europe by 2060 [30]. According to the landmark Framingham Heart Study in 1998, the mortality rates attributable to AF were 50% for men and 90% for women [31]. The Renfrew-Paisley Study in two Scottish towns in 2002 showed that AF increased all-cause mortality by 50% amongst men and 120% amongst women [32].

This work focuses on designing an arrhythmia beats detection method using single-lead ECG data with a patient-specific training model design since ECG waveforms vary significantly among different patients [3]. Furthermore, our arrhythmia classification work is based on the AAMI standards to make the proposed work comparable with representative related works. This work aims at achieving high classification (prediction) performance in the SVEB class and still meets the real-time ECG classification requirement. The proposed method uses Empirical Mode Decomposition (EMD) with resampling (EMDR), which resamples the first Intrinsic Mode Functions (*i.e.*, IMF 1) as a main input, for the proposed EMDR-LSTM architecture, and can improve the SVEB classification performance. By this method, we can achieve high classification performance in terms of accuracy, sensitivity, positive predictivity, and F1 and G scores for the SVEB class in all datasets (A, B, and C) [14]. Note that it is important to improve SVEB class detection because two subclasses in the SVEB class, which are Supraventricular Premature Beats (SPB) and Atrial Premature Beats (APB), can help assess if patients have high possibilities of AF/atrial flutter in the future [18, 19]. The SPB has also been shown to be one of the reasons to trigger AF [20].

The proposed work only uses single-lead ECG data instead of multilead ECG data. This is useful in some wearable health monitoring devices, where only one lead is available [2]. Using multilead ECG is more expensive in terms of time and computational resources for processing ECG signals since each lead needs to be processed. Multilead ECG also increases model complexity due to the increasing number of parameters in a model since using multilead ECG means more ECG signals, which makes the input length becomes longer. Single-lead ECG data are from Modified Limb lead II (MLII) because it is the most common lead that is used for a single lead record [2-4]. In addition, this work uses the LSTM as a classifier since the heartbeat activities are reflected in an ECG waveform, so there are temporal dependencies naturally existing in the ECG waveform [2]. The LSTM can capture such temporal dependencies in sequential data more efficiently compared to other types of neural networks [2].

1.1 Problem Statement

In a single lead ECG, detecting the SVEB class is considered more difficult than the other classes, such as normal class and VEB class [2, 3]. Related works [2-4] show low

classification performance, in terms of sensitivity, positive predictivity, and F1 and G scores, for detecting the SVEB class in a single-lead ECG. The low classification performance is related to the chosen features and the classification method. In this work, we focus on selecting a different combination of common ECG features, experimenting with a different resampling factor for the EMD feature, and also experimenting with a novel LSTM architecture for achieving high SVEB classification performance.

1.2 Contribution

For improving SVEB classification performance in a single-lead ECG, we proposed EMDR-LSTM that uses EMD with resampling (resamples the first IMF) for feature extraction and a novel LSTM architecture. The proposed EMDR-LSTM can improve and achieve the highest classification performance in terms of accuracy, sensitivity, positive predictivity, and F1 and G scores in the SVEB class, compared to representative related works [2-4]. To the best of our knowledge, the proposed EMDR-LSTM method is the first work that uses resampling the first IMF in the LSTM.

The rest of this paper is organized as follows. Section 2 describe the background and review related works. In Section 3, we present the proposed EMDR in a novel LSTM architecture. Section 4 shows experiment setup and experiment results. Finally, conclusions of this work and directions of future work are presented in Section 5.

2. BACKGROUND AND RELATED WORKS

This section describes the background on ECG, EMD and LSTM, and reviews representative related works on ECG signal classification techniques.

2.1 Background

2.1.1 R-peak and RR interval

As shown in Fig. 1, a typical waveform of a normal ECG consists of several points such as P, Q, R, S, T. Each of these points related to the heartbeat muscle movement and the way electrocardiogram records the heartbeat electrical impulse [28]. The R-peak position is the highest peak in the ECG waveform. In Fig. 2, an RR interval is the duration between two subsequent R-peak's, where the current R-peak position denotes as $R(i)$, the previous R-peak position denotes as $R(i-1)$, and the next R-peak position denotes as $R(i+1)$. The RR interval value is used as one of the concatenation materials for the input in this work.

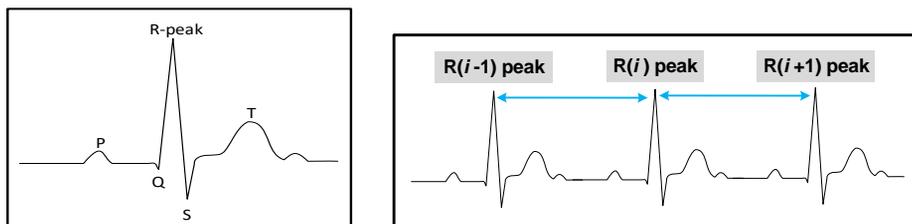


Fig. 1. A typical waveform of a normal ECG. Fig. 2. RR intervals in a normal ECG waveform.

2.1.2 EMD

The EMD is an adaptive time-space analysis method developed by N. E. Huang *et al.* [12]. Its effectiveness has already been demonstrated successful to many important problems encountered in processing biomedical, financial, geographical, and acoustic signals. The EMD has already been demonstrated to be very efficient for such ubiquitous tasks as denoising, time-frequency analysis [17]. The EMD decomposes a signal into a set of IMFs by a series of shifts. An IMF is a function that satisfies two conditions [35]:

1. In the whole data set, the number of extrema and the number of zero crossings must either equal or differ at most by one.
2. At any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero.

The shifting process terminates when the original signal contains no significant frequencies. The last IMF obtained is known as residue r .

2.1.3 LSTM architecture

The LSTM is a type of recurrent neural networks (RNNs) [16] used in the field of deep learning, and all RNNs have feedback loops in the recurrent layer. However, the training of traditional RNNs suffers from a vanishing gradient problem [35-37]. The LSTM addresses the problem via input, forget, and output gates [36] in a cell which regulate the information flow into and out of the cell to overcome the vanishing gradient problem of the traditional RNNs [37]. The LSTM can process single data points and sequences of data. We chose this architecture because this deep learning technique is suitable for handling time series data. Fig. 3 shows an LSTM architecture. The LSTM units includes cells, an input gate (I_t), an output gate (O_t), and a forget gate (F_t). A cell remembers values over arbitrary time intervals, and the other three gates regulate how the flow information comes in or out of the cell. It also includes an intermediate vector (M_t), the input vector at time t (X_t) and two state vectors, H_t and C_t which are carried from time $t - 1$ to time t .

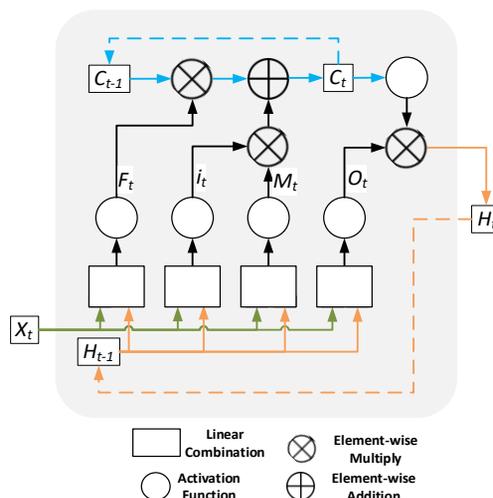


Fig. 3. Long short-term memory (LSTM) 1.

2.2 Related Works

Fig. 4 presents the classification techniques of ECG signals that compares the proposed EMDR-LSTM with representative related works. We categorize ECG signals into the single-lead and the multilead ECG. In the single-lead ECG, we split the category based on the AAMI standard and non-AAMI standard, which is also related to the dataset source and heart disease. Based on the most recent researches, we separate the classification methods into LSTM-based and 1-D CNN-based, and the proposed EMDR-LSTM is LSTM-based.

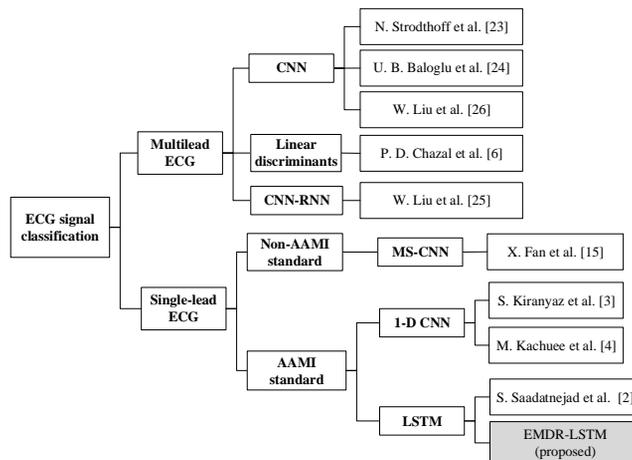


Fig. 4. Classification techniques of ECG signal.

2.2.1 Related works on multilead ECG and EMD

L. Y. Wang *et al.* [13] used EMD with singular value decomposition and the support vector machine (SVM) for classification. First, a signal is decomposed with EMD into a set of IMFs that are used to create a feature vector matrix. Then the feature vector matrix is decomposed using singular value decomposition and singular values of the matrix can be calculated as a feature of the ECG. After that, they used the SVM to identify the condition of arrhythmia. Izci *et al.* [8] used EMD for one of their main steps. Power spectral density (PS) and the variance of PS are extracted, which are applied to IMF 1 to IMF 7. Then classification using linear discriminant analysis is applied to differentiate normal and arrhythmic signals. Both related works [8, 13] were using the MIT-BIH arrhythmia database, However, they did not use the AAMI standards for their classification results, which makes their works not comparable to this proposed work.

Related works from [21, 22, 24] used CNNs to classify myocardial infarction beats into normal and abnormal classes. W. Liu *et al.* [23] used hybrid CNN-RNN for classifying the same disease using 12 lead ECG record data. However, their works [21-24] are not comparable to this proposed work since they used the multilead approach and did not focus on the heart disease. P. de Chazal *et al.* [6] used heartbeat intervals and morphology features from two ECG leads to classify ECG signals into classes based on the AAMI standards. They used a linear discriminant as the classifier. Although they used the patient-specific method and the same dataset source, their work used two ECG leads. Therefore, their work is also not comparable to the proposed work which uses only single-lead ECG.

2.2.2 Related works on single-lead ECG

There are related works on single-lead ECG. X. Fan *et al.* [13] proposed multi-scaled CNN (MS-CNN) to extract different feature sizes from ECG signal data. They used two inputs with two branches for their model with each branch implementing VGG architecture and used a different kernel size for each branch. Their work considered only two classes which are Atrial fibrillation (AF) and Normal and they used a dataset from Computing in Cardiology Challenge 2017 [5]. Since the dataset did not come from the MIT-BIH arrhythmia database and they did not use the AAMI standards, which makes their work not comparable to the proposed work.

S. Saadatnejad *et al.* [3] proposed an LSTM-based ECG classification for personal wearable devices. They used a combination of several input features based on the segmented ECG signal using the R-peak position, four distinct RR intervals and ECG signals that were transformed with discrete wavelet transform (DWT). Their model has three inputs. They used one or two layers RNN with LSTM cells and used multilayer perceptron (MLP) to produce classification results. They used the MIT-BIH arrhythmia database and implemented a patient-specific method to preprocess the dataset. They evaluated their work using a dual-lead ECG dataset, but also conducted evaluation on a single-lead ECG dataset using the MLII channel. Their method showed both lightweight and had good classification performance. In our proposed work, we used a new feature, which only resamples the first IMF with other common ECG features such as the ECG segment and the RR interval, and a novel LSTM architecture. The proposed work achieves higher SVEB classification performance in all metrics of single-lead ECG than this work, which will be shown later.

Kiranyaz *et al.* [3] used a one-dimensional (1-D) CNN to do the feature extraction and ECG classification based on the AAMI standards. Their method only utilizes a beat ECG segment and a beat-trio ECG segment, which both centered at the R-peak in the middle. Their method used only three layers of CNN and ran only 50 epochs with early stopping which can reduce the training time. In addition, their method is lightweight and have good classification performance. Compared to this work, the proposed work achieves higher SVEB classification performance in terms of accuracy, sensitivity, positive predictivity, and F1 and G scores. Kachuee *et al.* [4] used a 1-D residual CNN as their classification method. They also used the dataset from the MIT-BIH arrhythmia database and was based on the AAMI standards. Their work only used a segmented ECG as an input and used a 13-layer CNN architecture [4]. Compared to this work, the proposed work performs better in terms of SVEB classification performance and has less model complexity since the proposed work only utilizes a single layer LSTM architecture.

Table 1 summarizes the qualitative comparison of single-lead ECG signal classification approaches that includes the proposed work. First, the proposed work uses a unique feature that resamples only the first IMFs. Second, for the model architecture, we used a novel LSTM architecture, which is different from related works [2-4]. Besides, the proposed work achieves the highest performance results in terms of accuracy, sensitivity, positive predictivity, and F1 and G scores in the SVEB class and, compared to the representative related works [2-4]. While the preprocessing time of the proposed work is longer than that of related works [2-4], the inference time of the proposed work is lower than that of related works [2, 4]. Note that the quantitative evaluation of these single-lead ECG signal classification approaches will be shown in Section 4.

Table 1. Comparison of single-lead ECG signal classification approaches.

Approach	S. Saadatnejad <i>et al.</i> [1]	M. Kachuee <i>et al.</i> [3]	S. Kiranyaz <i>et al.</i> [2]	EMDR-LSTM (Proposed)
Dataset source	MIT-BIH arrhythmia	MIT-BIH arrhythmia	MIT-BIH arrhythmia	MIT-BIH arrhythmia
Classification classes	5 classes – AAMI standard	5 classes – AAMI standard	5 classes – AAMI standard	5 classes – AAMI standard
ECG lead	Dual lead or single-lead (MLII)	Single-lead (MLII)	Single-lead (MLII)	Single-lead (MLII)
Evaluation metric	Accuracy, sensitivity, specificity, positive predictivity, and F1 and G scores	Average accuracy	Accuracy, sensitivity, specificity, positive predictivity	Accuracy, sensitivity, specificity, positive predictivity, F1, G scores (highest in SVEB in all metrics except specificity)
Classification technique	LSTM (Two separate RNN-based models, model A and model B)	Deep residual CNN	1-D CNN	LSTM (One RNN-based model)
Features	RR-interval, ECG segment, DWT	Resampled ECG segment	A beat ECG Segment and a beat-trio ECG segment	EMD with resampling (resamples only the first IMF), RR interval, ECG segment
Preprocessing time	Short	Short	Short	Long
Inference time	Medium	High	Not available	Low

3. PROPOSED EMD WITH RESAMPLING IN A NOVEL LSTM ARCHITECTURE

3.1 Proposed EMD with Resampling for ECG Signal Preprocessing

In the proposed work, we used EMD with resampling (EMDR) as the main feature alongside other features as an input for the proposed LSTM architecture. As shown in Fig. 5, in the EMDR preprocessing design, we used an ECG signal as an input. This ECG signal consists of at least 10 R-peak's in order to extract the RR interval value. We used beat annotation that is in compliance with the AAMI standards to classify arrhythmia. As for the segmentation, the incoming digitized ECG samples are programmatically segmented into a sequence of heartbeats based on the R-peak position from the ground truth labeling so that each ECG segment has a fixed length of 0.25 seconds before R-peak and 0.45 seconds after R-peak 1. Note that there are R-peak position detection algorithms that are well established and highly accurate [42, 43]. This segmentation process also extracts the RR interval value for the current segmented ECG. In total, there are four RR interval values [2]. The first RR interval value is $rr1$, which comes from the distance between one previous R-peak position ($i - 1$) and the current R-peak position (i). The second RR interval value is $rr2$, which comes from distance from the current R-peak position (i) and one next R-peak position ($i + 1$). Third, $rr3$ is the average of 10 RR intervals from eight previous R-peak positions ($i - 8$) until one next R-peak position ($i + 1$). Originally, the $rr3$ feature by Saadatnejad [1] is from ($i - 4$) until ($i + 5$). In our pre-processing method, we modified the $rr3$ feature to reduce the prediction delay since the original one needs five next R-peak positions for classifying the current beat. In our modified $rr3$ feature, only one next R-peak

position is needed. Nevertheless, the modified $rr3$ feature does not affect the classification performance which has been validated by experiments. The fourth RR interval value is $rr4$, which is the average RR interval per patient (first 5 minutes). After getting all RR interval values, we concatenate it into one value ($rr1, rr2, rr3, rr4$), and it is denoted as RR_F . Then we apply normalization by dividing RR_F with the $rr4$ value [27]. RR_F will become one of the features used for an LSTM input.

There will be three outputs for our preprocessing method. We called these outputs as Preprocessing Outputs (POs). For the first PO, which PO1, the EMD function is applied to the segmented ECG (ECG_F). The EMD will decompose the ECG_F into several IMFs. In the proposed EMDR preprocessing technique, we only use the first IMF, and we only take the real number and remove the imaginary. After getting the first IMF, we apply it with a resampling function with a factor of 0.3. Note that we searched for the best resampling factor using a grid search and 0.3 is the best value for the resampling factor.

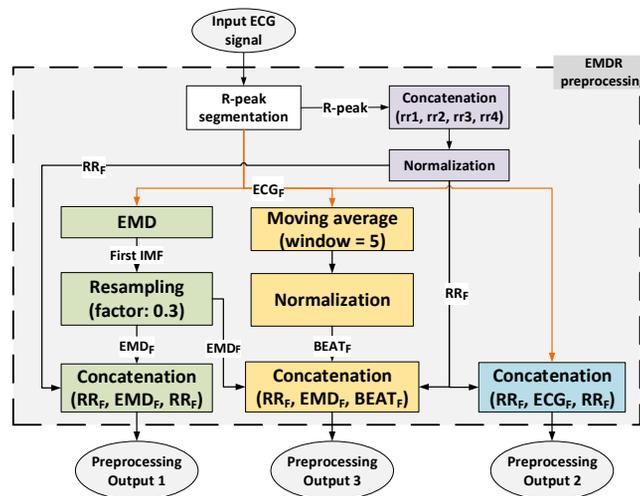


Fig. 5. EMDR preprocessing design.

Using EMD with resampling can enhance the SVEB classification performance in terms of Sen , and $F1$ and G scores since it helps extract the first IMF feature from segmented ECG for the SVEB class. Using resampling can reduce the complexity of the deep learning model by reducing the length of the first IMF by 70%, from the original EMD data length 250 reducing to 75, which reduces the number of parameters. This resampling the first IMF is denoted as EMD_F . After that, we concatenate it with the order of RR_F , EMD_F , and RR_F , to get the final Preprocessing Output 1 (PO1) results. For PO2, we directly use the ECG_F and RR_F with the following concatenation order, RR_F , ECG_F , and RR_F . In EMDR, by concatenating RR_F twice for PO1 and PO2, and ECG_F for PO2, the classification performance can be improved, which were our empirical research findings. This is for improving the classification performance by adding a raw ECG segment feature. For PO3, we do further preprocessing for ECG_F with downsampling using a moving average function with a window size of 5 and min-max normalization. The output for this process is called $BEAT_F$. This is to reduce the model complexity by reducing the input

length. The final PO3 result is a concatenation of RR_F , EMD_F , and $BEAT_F$. We employed a patient-specific training design method in the proposed work.

3.2 EMDR-LSTM Architecture

This proposed work is based on the LSTM architecture. We chose this architecture since it is suitable for handling the ECG waveform data. We cannot directly use related work [2]'s LSTM architecture, because our preprocessing method, EMDR, is different from that of related work [2]. This makes the three Pos' complexities and lengths of the proposed work be different, and several things are not optimized, such as LSTM hyperparameters and the number of layers for each LSTM input. Since the three POs are different, we need to design a novel LSTM architecture that is suitable for the POs to achieve the highest SVEB classification performance. In contrast to Saadatjad *et al.* [2] that uses two separate models with one or two LSTM layers for each input, we proposed a novel LSTM architecture that has one single model with a single LSTM layer only for each input and a single fully connected (FC) layer for achieving the highest SVEB classification performance.

Table 2 shows the LSTM architecture differences between the proposed EMDR-LSTM and Saadatjad *et al.* [1]. We chose a single model instead of two separate models used by Saadatjad *et al.* [1] because we found that our POs perform better in a single model instead of two. If we use two separate models, we need to have another FC layer and MLP to combine the result from each model. We found that an additional MLP layer makes the classification performance become worse compared to that using an FC only. We only use one LSTM layer instead of two, which also have an advantage of reducing the model complexity.

Table 2. The LSTM architecture differences between the proposed EMDR-LSTM and Saadatjad *et al.* [1].

	Saadatnejad et al. [2]	EMDR-LSTM (proposed)
RNN cells	LSTM	LSTM
Number of inputs	3	3
Number of models	two separate models	one single model
Number of RNN layers	1 or 2 layers	one layer only
Multilayer perceptron layers	Yes, 2 hidden layers	No, only a single fully connected layer

Fig. 6 shows how the proposed EMDR-LSTM works. The three outputs from EMDR, PO1, PO2, and PO3 are fed to the proposed LSTM architecture as inputs. The three POs will be reshaped to make them suitable for LSTM layer inputs, which are denoted as inputs 1, 2, and 3. Each input is fed and processed by an LSTM layer. The result from each LSTM layer is an array of LSTM features. Then the array of LSTM features from each input are concatenated into a large array of features. The final length of this concatenated array feature is the addition of a length of array features from each LSTM layer for inputs 1, 2, and 3. This concatenated array of LSTM features is then fed into the FC layer in order to get one-class classification results for the input beat.

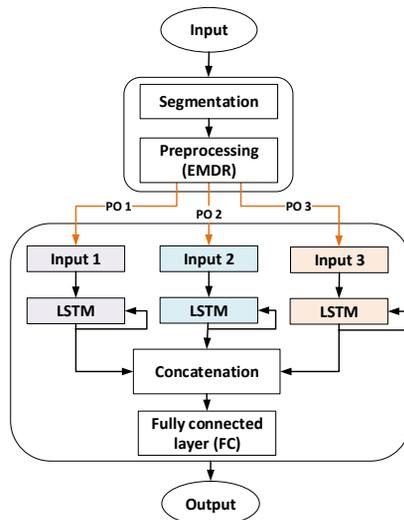


Fig. 6. EMDR-LSTM: Preprocessing and LSTM design architecture.

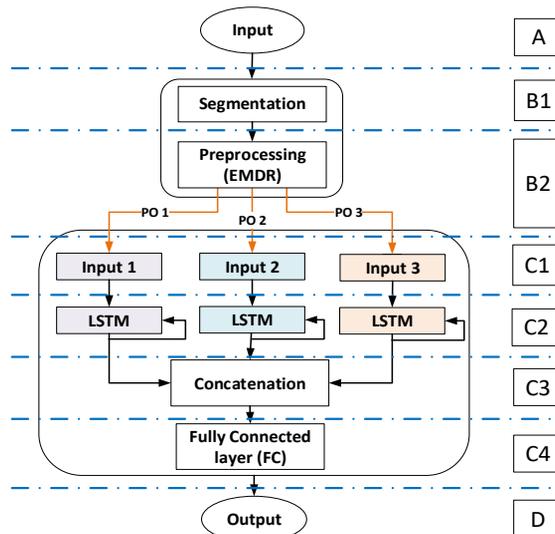


Fig. 7. EMDR-LSTM process flow.

3.3 EMDR-LSTM Process Flow

An example is used to illustrate each state of the proposed EMDR-LSTM. Fig. 7 shows the process flow of the EMDR-LSTM. Part A is the input, part B (B1 + B2) is for data preparation and processing, where part B1 is for ECG segmentation and part B2 is for EMDR-LSTM preprocessing. Part C (C1 + C2 + C3 + C4) is the LSTM architecture, where part C1 is for input reshaping for each LSTM, part C2 is the LSTM layer, part C3 is concatenation for the array features results from the LSTM layer for each input, and part C4 is the FC layer. The last part is part D which is the output for the ECG classification prediction result.

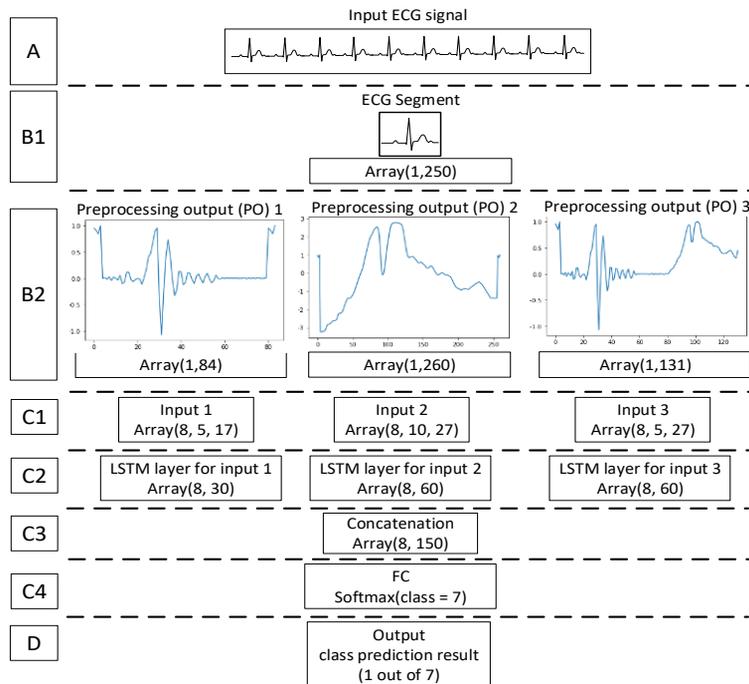


Fig. 8. Intermediate results of the EMDR-LSTM process flow.

Fig. 8 shows the intermediate results of the EMDR-LSTM process flow and the detailed data shape for each part. Part A is an input ECG signal, with at least 10 R-peak's, which consists of eight previous R-peak's, one current R-peak, and one next R-peak positions. Part B1 is the process of R-peak detection and R-peak based beat segmentation, where the results are segmented ECG and the RR interval values, as explained in Section 3.1. In part B2, after acquiring the ECG segment and RR interval values, the preprocessing is started, where it applies the EMD with resampling to the ECG segment and proceeds the preprocessing method, as explained in Section 3.1. There are three POs, which are PO1, PO2, and PO3 from the results of part B2. PO1 is the preprocessed signal with shape (1, 84), PO2 has shape (1, 260) and PO3 has shape (1,131).

Each PO will then be reshaped in order to make it suitable for an LSTM input, which is done in part C1: the final shape (8, 5, 17) for input 1, (8, 10, 27) for input 2 and (8, 5, 27) input 3. Each input shape is related to the number of batches, which is 8 and the timesteps, which are 5, 10, and 5 for inputs 1, 2, and 3, respectively, in the LSTM hyperparameters and each PO length settings. Part C2 shows the LSTM layer for each input with each array feature results from the LSTM layer. The output array feature is related to the neuron setting for each LSTM layer, which are 30, 60, and 60 for inputs 1, 2, and 3, respectively. In part C3, the array of features from each LSTM layer is concatenated into one large array of features, which makes the length become 150 because it is a concatenation of the array features from the LSTM layer for each input. Part C4 uses the concatenated array of features from part C3 as an input for the FC

layer which applies the *Softmax* function in order to classify into 7 classes. Finally, part D is the output from the FC layer, which is the classification prediction result of one final class for the input beat, 1 predicted class out of 7 available classes.

4. EXPERIMENT SETUP AND RESULTS

In the proposed LSTM architecture, the setting of hyperparameters is an important matter. We used Neuron (N) settings $N1 = 30$ for the LSTM layer for input 1, and $N2 = 60$ for input 2, and $N3 = 60$ for input 3. We obtained optimum neuron values by using the grid search from 10 – 100 with a step value of 10 for each increment. The proposed architecture was trained with epoch 110 and the number of batches per epoch was set to 8. The values of timesteps for LSTM inputs is 5, 10, and 5 for inputs 1, 2, and 3, respectively. We employed early stopping to prevent overfitting and to reduce training time. We used two criteria in early stopping: first, if the training loss is below 0.11 and second, if the training accuracy is above 0.98. The proposed LSTM architecture will stop the training to avoid overfitting if either of these criteria is met. We used the categorical cross-entropy function [38] as the loss function and the training accuracy, in the range from 0 to 1, is defined as the number of correct predictions divided by the number of total predictions during the training process [39]. For the FC layer, we used the *softmax* activation function to take the classification result with the highest probability to be the final result for the beats. Other parameters for LSTM cells were the *tanh* activation function and *sigmoid* for recurrent activation. For kernel initialization, we used Glorot uniform, and recurrent initialization was orthogonal with gain value 1.0 and biases were initialized to zeros. The random seed was set to 18 for kernel initialization in each layer. We chose the trained LSTM model with the one that has the minimum training loss value.

4.1 Dataset and Tools

The dataset from the MIT-BIH ECG arrhythmia database [14] was used for evaluating the proposed EMDR-LSTM, and we chose this dataset to make our evaluation comparable with that of related works [2-4]. This dataset consists of 48 records from 47 patients. Each record has two leads; the first one is MLII and the second one is modified lead V1 or V2, V4, and V5. Note that for beat annotations, two or more cardiologists independently annotated each record in the MIT-BIH dataset [14]. This dataset contains two sets of data based on patient *id*. It is called DS100 with an *id* range from 100-199 and DS200 from 200-299. This dataset includes data for both normal and abnormal beats. Based on the AAMI standards [7], patient *id* with paced beats (102, 104, 107, and 217) are excluded from testing and training datasets. This work uses a single-lead ECG approach that uses the first lead, the MLII, since it is the most common lead for a single-lead experiment.

Hardware specifications for training the LSTM model were Intel i9 CPU, NVIDIA GeForce RTX 2080 SUPER with 8 gigabytes GPU memory, and 32 gigabytes RAM memory. The LSTM architecture was implemented in Keras deep learning API [10] that ran on the top of TensorFlow [9]. The EMD algorithm was implemented in Python that used the EMD-signal library [11]. We used the Python 3.0 scripting language for data pre-processing and for developing of the LSTM model.

4.2 Patient-Specific Training

We used a patient-specific training method in the proposed EMDR-LSTM architecture to make it comparable to representative related works [2-4]. Patient-specific training means that each patient has his/her patient-specific training data and his/her model is trained using his/her data. Each patient-specific training data come from the first 5 minutes of each patient data, and are combined with global data. The global data come from patient DS100. We randomly added the global data to each first 5 minutes of patient data with the size of the maximal 30 global data for each class representative. For example, for patient *id* 200, we added a maximal of 30 beat data from the global data (random chosen) for each class representative which is 7, to the first 5 minutes of patient *id* 200 data, so the maximal global data that used were $30 \times 7 = 210$. The random global data are the same for all patients.

For test data, the first 5 minutes of each patient data were removed. Since we employed the patient-specific training method, each patient has his/her model instead of one model for all patients. Fig. 9 shows how we combine the data and how the model is generated for each patient.

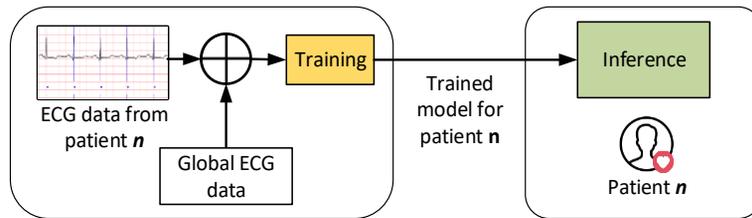


Fig. 9. Patient-specific training and ECG signal classification [2].

4.3 Evaluation Metrics

The proposed EMDR-LSTM was evaluated and compared with representative related works [2-4] in terms of accuracy (Acc), sensitivity (Sen), specificity (Spe), positive predictivity (Ppr), and $F1$ and G scores. This work performs the multi-class classification and evaluates the VEB and SVEB classes in binary classification. The terms TP, TN, FP, and FN denote as true positive, true negative, false positive, and false negative, respectively. The performance metrics are summarized as follows:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Sen = \frac{TP}{TP + FN} \quad (2)$$

$$Spe = \frac{TN}{TN + FP} \quad (3)$$

$$Ppr = \frac{TP}{TP + FP} \quad (4)$$

$$F1 = \frac{2}{\frac{1}{Sen} + \frac{1}{Ppr}} \quad (5)$$

$$G = \sqrt{Sen \times Ppr} \tag{6}$$

We classify the results into seven arrhythmia classes where we split the normal class (N) into two conduction abnormalities, which are left bundle branch block (L) and right bundle branch block (R) to increase resolution [2]. Then the seven classes were combined into five classes based on the AAMI standards to make the proposed work comparable to related works [2-4]. Note that five arrhythmia classes based on the AAMI standards are Normal (N), Supraventricular ectopic beats (SVEB or S), Ventricular ectopic beats (VEB or V), Fusion (F), and Unknown (Q). Table 3 shows the seven classes and its relation to the five classes that is based on the AAMI standards. There is a specific dataset of patient *id*'s for testing purposes based on related work [1]. We follow this dataset to make the proposed work comparable to related works [1-3]. The datasets are dataset A: VEB: 200, 202, 210, 213, 214, 219, 221, 228, 231, 233 and 234 (11 Patient *id*'s), SVEB: the same as VEB, plus 212, 222, and 232 (14 Patient *id*'s), dataset B: 100, 103, 105, 111, 113, 117, 121, 123, 200, 202, 210, 212, 213, 214, 219, 221, 222, 228, 231, 232, 233, 234 (22 patient *id*), and dataset C: 200, 201, 202, 203, 205, 207, 208, 209, 210, 212, 213, 214, 215, 219, 220, 221, 222, 223, 228, 230, 231, 232, 233, and 234, (24 patient *id*'s).

In addition, we derived the inference time for each model. Note that inference time, which is the latency (or the response time) of an AI model, is an important metric, especially for real-time or time-critical applications, such as an ECG medical monitoring application, which is the target application of this paper. We computed inference time using the *timeit* library in Python and ran on Google Colab. We used a patient-specific method to calculate the inference time of each model. The testing data consist of 44 patients, denoted as *Total Patients* (TPs), and a total of 83714 segmented beats, denoted as *Total Segmented Beats* (TSBs). First, we calculate *Inference Time Per Patient* (ITPP) and summarize all the patient's inference time to get *Total Inference Time* (TIT). Then we calculate the *average of Inference Time per Patient* (aITP) and the *average of Inference Time per Beat* (aITB). We also calculate the *Longest Inference Time from All Patients* (LITAP) and the *Shortest Inference Time from All Patients* (SITAP), as follows:

$$aITP = \frac{TIT}{TPs}, \tag{7}$$

$$aITB = \frac{TIT}{TSBs}, \tag{8}$$

$$LITB = \max((ITPP_1), \dots, (ITPP_n)) \tag{9}$$

$$SITAP = \min((ITPP_1), \dots, (ITPP_n)) \tag{10}$$

Table 3. Heartbeat classes [2, 7].

5 Labels	7 Labels	Hearbeat Types
N	N	Normal beat, atrial escape beat, junctional escape beat
	L	Left bundle branch block beat
	R	Right bundle branch block beat
S	S	Atrial premature beat, aberrated atrial premature beat, junctional premature beat, supraventricular premature beat
V	V	Premature ventricular contraction, ventricular escape beat
F	F	Fusion of ventricular and normal beat
Q	Q	Paced beat, fusion of paced and normal beat, unclassifiable beat

4.4 Evaluation Results

4.4.1 Evaluation of ECG signal classification

For evaluation, we implemented the proposed model so as to compare with representative related works [2-4]. Note that S. Saadatnejad *et al.* [2] provides online source codes for data and their model approach, but it is only for dual-lead ECG. To produce the results of related work [2] in terms of a single lead under the same environment, we recreated the preprocessing and a deep learning model based on their work. We also recreated the work by Kachuee *et al.* [4] and implemented their work in a patient-specific training way. We make sure our implementation is correct by comparing and checking the similarities between our produced results to the results of their works. As to Kiranyaz *et al.* [3], based on our observation, the total beats data difference between the two approaches is only 69. Therefore, we directly used their results because there is no significant results difference if the total beats data difference is not more than 250. Since the training of deep learning depends on randomness such as initial weights and the optimizer used, we ran the test of the proposed work 10 times and took the average. Tables 4 and 5 show the classification results comparisons between the proposed model and related works [2-4] for VEB and SVEB classes, respectively. In Table 4, the proposed EMDR-LSTM model achieves better classification performance compared to related works [2-4] in VEB class, for datasets A and B. However, in Table 4, the proposed model achieves lower classification performance compared to related works [2, 3] in VEB class, for dataset C. This is because we designed the preprocessing method and fine-tuned the LSTM hyperparameter to enhance the SVEB class instead of the VEB class. In Table 5, the proposed model has the highest SVEB classification performance in terms of *Acc*, *Sen*, *Ppr*, and *F1* and *G* scores compared to all related works [2-4] in all datasets. In addition, the proposed model has the lowest standard deviations compared to related works [2, 4] in both VEB and SVEB classes, as shown in Tables 4 and 5, respectively, which indicates the stability of the proposed model. By utilizing two early stopping criteria [40] that were not used by related works [2, 4] and by using mini-batch gradient descent [41] with a smaller mini-batch size than that of related work [2] for training, the proposed model can avoid the overfitting problem and can thus produce the model with the lowest standard deviations.

In addition, we derived the total number of parameters (*Total params*) and MACs (multiply-accumulate operations) to reflect cost, and inference time, which is an important performance metric for real-time or time-critical applications, to reflect an additional performance measure besides the original performance measures (*Acc*, *Sen*, *Spe*, *etc.*) that have been derived. Based on the experiment results in Table 6, the *Total params* in the proposed EMDR-LSTM model is higher than that of Saadatnejad *et al.* [2], but lower than that of Kachuee *et al.* [4]. As to the MACs, the proposed model is higher than the other two related works. However, Table 7 shows that the proposed model has the lowest inference time in terms of LITAP, SITAP, aITP and aITB, compared to related works [2, 4]. Note that we repeated this experiment 10 times and took the average for each of the above performance metrics. The reasons that the proposed model has the lowest inference time among the three models are described as follows. Saadatnejad *et al.* [2] uses two RNN-based models, named model *alpha* and model *beta*. Both models use an RNN with LSTM cells with one or two layers, where model *alpha* consists of *two branches* while model *beta* has only *one branch*. Kachuee *et al.* [4] uses a model consists of five residual blocks and each residual block contains two convolutional (CNN) layers in

series. In total, the resulting network is a *deep network consisting of 13-layer CNN architecture* [4]. In contrast, the proposed EMDR-LSTM model uses an LSTM architecture that has *three LSTM layers branches* to process preprocessing outputs PO1, PO2, and PO3, in parallel. This indicates that the proposed model makes better use of underlying parallel software/hardware, where the three LSTM layers branches use preprocessing outputs PO1, PO2, and PO3, in parallel, as shown in Fig. 7, in contrast to S. Saadatnejad *et al.* [2] and Kachuee *et al.* [4]. In summary, although the proposed EMDR-LSTM has higher preprocessing cost and higher computational complexity in terms of MACs, it not only has higher overall classification performance, but also has lower standard deviations of all the performance metrics and lower inference time, which are important performance metrics for real time or time-critical applications, *e.g.*, ECG medical monitoring applications, compared to representative related works [2, 4].

Table 4. Comparison of single-lead ECG signal classification performance for VEB class in terms of Average and standard deviation of each performance metric.

Dataset	Approach	VEB					
		<i>Acc</i>	<i>Sen</i>	<i>Spe</i>	<i>Ppr</i>	<i>Fl</i>	<i>G</i>
A	Saadatnejad <i>et al.</i> [2]	98.86 ± 0.18	94.69 ± 0.49	99.42 ± 0.19	95.62 ± 1.34	95.15 ± 0.72	95.15 ± 0.73
	Kachuee <i>et al.</i> [4]	97.60 ± 0.99	91.95 ± 1.71	98.36 ± 1.18	88.27 ± 7.07	90.07 ± 3.55	90.09 ± 3.43
	Kiranyaz <i>et al.</i> [3]*	98.90	95.90	99.40	96.20	96.05	96.05
	EMDR-LSTM (proposed)	99.26 ± 0.06	95.60 ± 0.56	99.75 ± 0.08	98.09 ± 0.59	96.83 ± 0.24	96.84 ± 0.24
B	Saadatnejad <i>et al.</i> [2]	99.30 ± 0.10	94.36 ± 0.53	99.64 ± 0.10	94.76 ± 1.33	94.56 ± 0.71	94.56 ± 0.71
	Kachuee <i>et al.</i> [4]	98.27 ± 0.49	91.57 ± 1.70	98.74 ± 0.48	83.40 ± 5.24	87.29 ± 3.21	87.39 ± 3.10
	EMDR-LSTM (proposed)	99.41 ± 0.07	95.42 ± 0.55	99.68 ± 0.09	95.44 ± 1.21	95.43 ± 0.53	95.43 ± 0.53
C	Saadatnejad <i>et al.</i> [2]	98.14 ± 0.14	85.27 ± 0.73	99.53 ± 0.15	95.07 ± 1.45	89.90 ± 0.68	90.04 ± 0.71
	Kachuee <i>et al.</i> [4]	97.04 ± 0.40	83.32 ± 1.82	98.51 ± 0.34	85.75 ± 2.94	84.52 ± 2.02	84.53 ± 2.03
	Kiranyaz <i>et al.</i> [3]*	98.60	95.00	98.10	89.50	92.17	92.21
	EMDR-LSTM (proposed)	98.17 ± 0.10	87.36 ± 1.02	99.33 ± 0.09	93.34 ± 0.79	90.25 ± 0.50	90.30 ± 0.49

* Since the source code of Kiranyaz *et al.* [3] is not available, the average of each metric is directly coming from [3].

Table 5. Comparison of single-lead ECG signal classification results for SVEB class in terms of Average and standard deviation of each performance metric.

Dataset	Approach	SVEB					
		<i>Acc</i>	<i>Sen</i>	<i>Spe</i>	<i>Ppr</i>	<i>Fl</i>	<i>G</i>
A	Saadatnejad <i>et al.</i> [2]	96.7 ± 0.25	80.90 ± 1.78	97.81 ± 0.35	68.41 ± 3.13	74.13 ± 1.30	74.39 ± 1.16
	Kachuee <i>et al.</i> [4]	95.90 ± 0.94	80.38 ± 3.34	96.81 ± 1.04	59.61 ± 7.05	68.45 ± 4.69	69.22 ± 4.14
	Kiranyaz <i>et al.</i> [3]*	96.40	68.80	99.50	79.20	73.63	73.82
	EMDR-LSTM (proposed)	98.07 ± 0.21	83.44 ± 0.79	98.93 ± 0.23	82.10 ± 3.06	82.76 ± 1.52	82.77 ± 1.50
B	Saadatnejad <i>et al.</i> [2]	97.77 ± 0.17	80.94 ± 1.74	98.45 ± 0.22	67.74 ± 2.85	73.75 ± 1.21	74.05 ± 1.07
	Kachuee <i>et al.</i> [4]	96.96 ± 0.73	80.02 ± 3.20	97.64 ± 0.78	57.78 ± 7.35	67.11 ± 5.10	68.00 ± 4.49
	EMDR-LSTM (proposed)	98.57 ± 0.15	83.35 ± 0.77	99.18 ± 0.16	80.42 ± 3.01	81.86 ± 1.56	81.87 ± 1.52
C	Saadatnejad <i>et al.</i> [2]	96.94 ± 0.18	62.97 ± 1.58	98.61 ± 0.23	69.16 ± 3.20	65.92 ± 1.16	65.99 ± 1.22
	Kachuee <i>et al.</i> [4]	95.62 ± 0.80	60.37 ± 2.58	97.36 ± 0.89	53.02 ± 7.84	56.46 ± 4.17	56.58 ± 3.93
	Kiranyaz <i>et al.</i> [3]*	96.40	64.60	98.60	62.10	63.33	63.34
	EMDR-LSTM (proposed)	97.64 ± 0.14	65.63 ± 1.12	99.22 ± 0.15	80.67 ± 2.83	72.38 ± 1.34	72.76 ± 1.43

* Since the source code of Kiranyaz *et al.* [3] is not available, the average of each metric is directly coming from [3].

Table 6. Comparison of Total number of parameters and MACs.

Approach	Saadatnejad <i>et al.</i> [2]	Kachuee <i>et al.</i> [4]	EMRD-LSTM (Proposed)
Total params	33,751	53,957	49,057
MACs	1,256,760	3,586,559	3,812,250

Table 7. Comparison of average inference time in terms of LITAP, SITAP, aITP and aITPB.

Approach	LITAP (Sec)	SITAP (Sec)	aITP (Sec)	aITB (Sec)
Saadatnejad <i>et al.</i> [2]	0.51604	0.20940	0.29751	0.00016
Kachuee <i>et al.</i> [4]	0.81132	0.32479	0.51333	0.00027
EMDR-LSTM (Proposed)	0.33155	0.13293	0.19832	0.00010

4.4.2 Evaluation of EMD and resampling features

We also conducted ablation studies to know which feature has a significant effect on the proposed method. We did three kinds of evaluation. The first is without EMD and without resampling, which means ECG segment beats are processed directly without preprocessing. The second is with EMD and without resampling, which checks the effect of resampling that is applied to the EMD. The third is with EMD and with resampling, which is the proposed method.

In Table 8, for the VEB class, the resamples first IMF feature does not affect the results significantly. In Table 9, it shows the proposed preprocessing method which uses the EMD with resampling achieves the highest performance results in terms of *Sen*, and *F1* and *G* scores in all the datasets (A, B, C) in the SVEB class compared to that if we did not use the EMD or the resampling method. In Table 9 the proposed resampled first IMFs feature shows to have a significant contribution to the SVEB class, since if the feature is removed, *Sen*, and *F1* and *G* scores decrease significantly. The resampling method that is applied to the first IMF contributes to increase *Sen*, although it decreases *Ppr* and *Spe*. We chose the one with the highest *Sen*, *F1*, and *G* scores results since in the medical field sensitivity (*Sen*) is an important metric for evaluation. Another advantage of using the resampling method is that we can reduce the LSTM model complexity by reducing the number of parameters since the input length becomes short.

Table 8. Ablation studies of the main EMD features in the VEB class.

Dataset	EMD Features	VEB					
		<i>Acc</i>	<i>Sen</i>	<i>Spe</i>	<i>Ppr</i>	<i>F1</i>	<i>G</i>
A	w/o EMD and w/o resampling	99.15	95.17	99.68	97.56	96.35	96.36
	With EMD, w/o resampling	99.14	95.52	99.63	97.18	96.34	96.35
	EMD with resampling (proposed)	99.26	95.60	99.75	98.09	96.83	96.84
B	w/o EMD and w/o resampling	99.32	94.93	99.63	94.65	94.79	94.79
	With EMD, w/o resampling	99.34	95.37	99.61	94.42	94.89	94.89
	EMD with resampling (proposed)	99.41	95.42	99.68	95.44	95.43	95.43
C	w/o EMD and w/o resampling	98.19	86.11	99.49	94.80	90.25	90.35
	With EMD, w/o resampling	97.96	85.83	99.26	92.60	89.09	89.15
	EMD with resampling (proposed)	98.17	87.36	99.33	93.34	90.25	90.30

Table 9. Ablation studies of the main EMD features in the SVEB class.

Dataset	EMD Features	SVEB					
		<i>Acc</i>	<i>Sen</i>	<i>Spe</i>	<i>Ppr</i>	<i>F1</i>	<i>G</i>
A	w/o EMD and w/o resampling	97.80	77.23	99.01	81.99	79.54	79.57
	With EMD, w/o resampling	98.12	79.56	99.21	85.44	82.40	82.45
	EMD with resampling (proposed)	98.07	83.44	98.93	82.10	82.76	82.77
B	w/o EMD and w/o resampling	98.20	76.48	99.07	76.89	76.68	76.68
	With EMD, w/o resampling	98.61	79.44	99.38	83.88	81.60	81.63
	EMD with resampling (proposed)	98.57	83.35	99.18	80.42	81.86	81.87
C	w/o EMD and w/o resampling	97.27	57.99	99.21	78.44	66.68	67.44
	With EMD, w/o resampling	97.50	61.58	99.28	80.83	69.90	70.55
	EMD with resampling (proposed)	97.64	65.63	99.22	80.67	72.38	72.76

We did experiments on using different resampling factor values. We conducted a grid search from 2 to 5 with a step value of 1. In the VEB class, the resampling factor does not significantly affect the classification performance. On the other hand, in the SVEB class, it has a significant effect on the classification performance. Fig. 10 is the evaluation results of using dataset A in the SVEB class, which shows the resampling factor 0.3 achieving the highest *Ppr*, and *F1* and *G* scores and achieving the second highest result on *Sen* compared to the other resampling factors. Evaluation of datasets B and C in the SVEB class shows similar improvements of the same metrics as those of dataset A, as shown in Figs. 11 and 12, respectively. The evaluation results show that resampling factor 0.2 achieves the highest *Sen* and the resampling factor 0.3 achieves the second highest *Sen*. However, we chose the resampling factor 0.3 as the proposed resampling factor because with this value it can achieve the highest classification performance in terms of *Ppr*, and *F1* and *G* scores in the SVEB class, among all resampling factors.

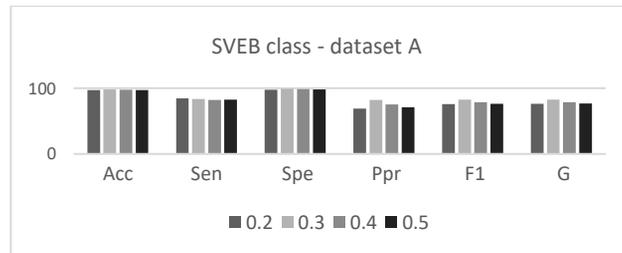


Fig. 10. Resampling factor evaluation in SVEB class – dataset A.

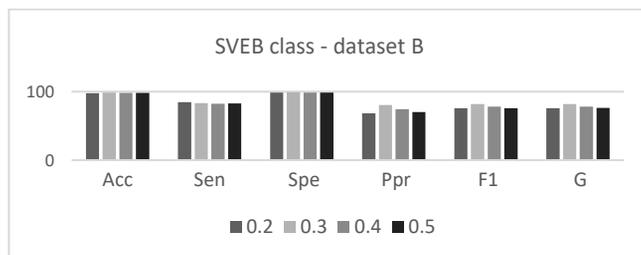


Fig. 11. Resampling factor evaluation in SVEB class – dataset B.

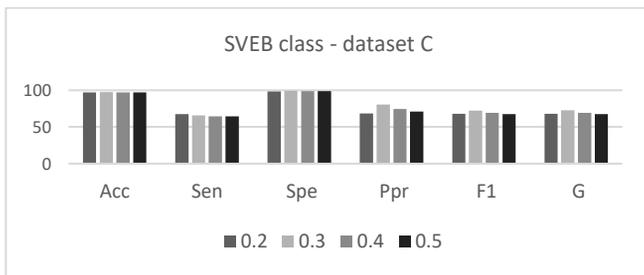


Fig. 12. Resampling factor evaluation in SVEB class – dataset C.

4.4.3 Evaluation of preprocessing time

We evaluated the preprocessing time of the proposed method compare to representative related works [2-4]. As shown in Fig. 13, the preprocessing time of the EMDR-LSTM is longer, compared to related works [2-4]. It is because the EMD takes time to decompose signals into a set of IMFs, even we only use the first IMF. Although the preprocessing time is longer, the EDMR-LSTM is still able to meet the real-time ECG classification requirement. This is because that the maximum heartbeat per minute for a person is around 220 beats, and our preprocessing time for each beat is only 7 milliseconds.

4.4.4 Discussion

Our work investigates the arrhythmias that are not imminently life-threatening but may require therapy to prevent further problems [6]. We plan to integrate our work to wearable sensor devices [33, 34] for monitoring and notifying a user to have an arrhythmia survey if any abnormality is detected. For example, if a user has obvious structural heart diseases and has more than 40 atrial premature *beats/h*, there is a very high possibility of AF/atrial flutter development in the next 2 to 3 years [44]. With this integration, the user can be advised for an intensive arrhythmia survey since he/she may have a high possibility to develop into AF/atrial flutter in the future.

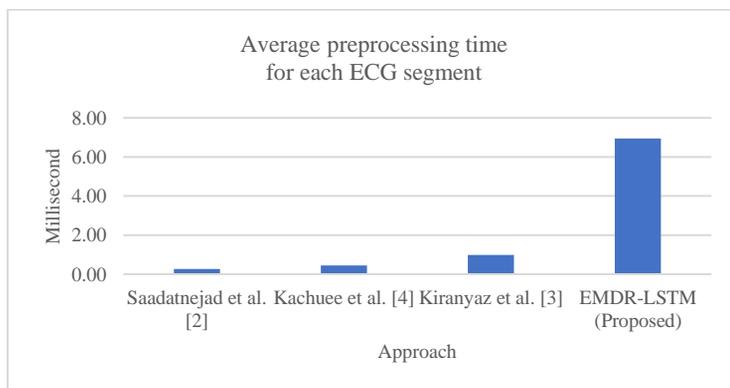


Fig. 13. Comparison of average preprocessing time for each ECG segment.

5. CONCLUSION AND FUTURE WORK

5.1 Conclusion

In this paper, we have presented a novel EMDR-LSTM architecture to enhance SVEB classification performance. The EMDR-LSTM has been proven to achieve the highest SVEB classification performance in terms of accuracy, sensitivity, positive predictivity, and F1, and G scores in all datasets (A, B, and C) compared to the three representative related works [2-4]. The EMDR-LSTM achieves 98.07%, 83.44%, 82.10%, 82.76%, 82.77% in dataset A, 98.57%, 83.35%, 80.42%, 81.86%, 81.87% in dataset B and 97.64%, 65.63%, 80.67%, 72.38%, 72.76%, in dataset C in terms of accuracy, sensitivity, positive predictivity, F1, and G scores, respectively, in SVEB class. The proposed EMD with resampling method that resamples the first IMF has been shown to enhance the SVEB classification performance in terms of sensitivity, and F1 and G scores for all datasets, compared to the one without using the EMD or the resampling method. However, for EMD, since it takes time to decompose the signal into a set of IMFs, this makes preprocessing time long. Nevertheless, it still meets the real-time ECG classification requirement. In addition, although the proposed EMDR-LSTM has higher computational complexity in terms of MACs, it has lower standard deviation of the performance metrics and lower inference time, which are important performance metrics for real time or time-critical applications, e.g., ECG medical monitoring applications, compared to the representative related works [2, 4].

5.2 Future Work

In future work, we will evaluate the EMDR-LSTM for classifying live data from common wearable ECG monitoring devices, such as Holter monitor [33, 34], which can be used to notify a user to have a heart health check if an abnormal condition is detected. We will also evaluate the EMDR-LSTM using multilead data to see if it is also applicable to handle multilead data, such as data for myocardial infarction. In addition, we will experiment on different deep learning architectures, such as a hybrid CNN-RNN, to see if it can enhance classification performance since the CNN and RNN have been shown to have good classification performance in the VEB and SVEB classes, respectively.

REFERENCES

1. H. H. Publishing, "Atrial fibrillation: Common, serious, treatable," <https://www.health.harvard.edu/heart-health/atrial-fibrillation-common-serious-treatable>, 2020.
2. S. Saadatnejad, M. Oveisi, and M. Hashemi, "LSTM-based ECG classification for continuous monitoring on personal wearable devices," *IEEE Journal of Biomedical and Health Informatics*, Vol. 24, 2020, pp. 515-523.
3. S. Kiranyaz, T. Ince, and M. Gabbouj, "Real-time patient-specific ECG classification by 1-D convolutional neural networks," *IEEE Transactions on Biomedical Engineering*, Vol. 63, 2016, pp. 664-675.
4. M. Kachuee, S. Fazeli, and M. Sarrafzadeh, "ECG heartbeat classification: A deep transferable representation," in *Proceedings of IEEE International Conference on Healthcare Informatics*, 2018, pp. 443-444.

5. A. L. Goldberger, L. A. N. Amaral, L. Glass, *et al.*, "PhysioBank, PhysioToolkit, and PhysioNet," *Circulation*, Vol. 101, 2000, No. E215-20.
6. P. Dechazal, M. O'dwyer, and R. Reilly, "Automatic classification of heartbeats using ECG morphology and heartbeat interval features," *IEEE Transactions on Biomedical Engineering*, Vol. 51, 2004, pp. 1196-1206.
7. *AAMI-Recommended Practice: Testing and Reporting Performance Results of Ventricular Arrhythmia Detection Algorithms*, Association for the Advancement of Medical Instrumentation, VA, 1987.
8. E. Izci, M. A. Ozdemir, R. Sadighzadeh, and A. Akan, "Arrhythmia detection on ECG signals by using empirical mode decomposition," *Medical Technologies National Congress*, 2018, pp. 1-4.
9. M. Abadi *et al.*, "TensorFlow: A system for large-scale machine learning," in *Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation*, 2016, pp. 265-283.
10. Keras-Team, "keras-team/keras," <https://github.com/keras-team/keras>, 2020.
11. Laszukdawid, "laszukdawid/PyEMD," <https://github.com/laszukdawid/PyEMD>, 2020.
12. N. E. Huang, Z. Shen, S. R. Long, *et al.*, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," in *Proceedings of the Royal Society of London, Series A: Mathematical, Physical and Engineering Sciences*, Vol. 454, 1998, pp. 903-995.
13. Y.-J. Wang, L.-X. Song, and S.-Q. Kang, "Arrhythmia recognition based on EMD and support vector machines," in *Proceedings of the 4th International Conference on Bioinformatics and Biomedical Engineering*, 2010, pp. 1-4.
14. G. Moody and R. Mark, "The impact of the MIT-BIH arrhythmia database," *IEEE Engineering in Medicine and Biology Magazine*, Vol. 20, 2001, pp. 45-50.
15. X. Fan, Q. Yao, Y. Cai, *et al.*, "Multiscaled fusion of deep convolutional neural networks for screening atrial fibrillation from single lead short ECG recordings," *IEEE Journal of Biomedical and Health Informatics*, Vol. 22, 2018, pp. 1744-1753.
16. S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, Vol. 9, 1997, pp. 1735-1780.
17. D. M. Klionskiy, D. I. Kaplun, and V. V. Geppener, "Empirical mode decomposition for signal preprocessing and classification of intrinsic mode functions," *Pattern Recognition and Image Analysis*, Vol. 28, 2018, pp. 122-132.
18. S. Kochhäuser, D. G. Dechering, R. Dittrich, *et al.*, "Supraventricular premature beats and short atrial runs predict atrial fibrillation in continuously monitored patients with cryptogenic stroke," *Stroke*, Vol. 45, 2014, pp. 884-886.
19. M. Martínez-Sellés, I. F. Lozano, A. Baranchuk, *et al.*, "Should we anticoagulate patients at high risk of atrial fibrillation?" *Revista Española de Cardiología*, (English Editing), Vol. 69, 2016, pp. 374-376.
20. S. Nortamo, T. V. Kenttä, O. Ukkola, *et al.*, "Supraventricular premature beats and risk of new-onset atrial fibrillation in coronary artery disease," *Journal of Cardiovascular Electrophysiology*, Vol. 28, 2017, pp. 1269-1274.
21. C. Ferreira, R. Providência, M. J. Ferreira, and L. M. Gonçalves, "Atrial fibrillation and non-cardiovascular diseases: A systematic review," *Arquivos Brasileiros de Cardiologia*, Vol. 105, 2015, pp. 519-526.

22. T. M. Munger, L. Q. Wu, and W. K. Shen, "Atrial fibrillation," *Journal of Biomedical Research*, Vol. 28, 2014, pp. 1-17.
23. N. Strodthoff and C. Strodthoff, "Detecting and interpreting myocardial infarction using fully convolutional neural networks," *Physiological Measurement*, Vol. 40, 2019, p. 015001.
24. U. B. Baloglu, M. Talo, O. Yildirim, R. S. Tan, and U. R. Acharya, "Classification of myocardial infarction with multi-lead ECG signals and deep CNN," *Pattern Recognition Letters*, Vol. 122, 2019, pp. 23-30.
25. W. Liu, F. Wang, Q. Huang, S. Chang, H. Wang, and J. He, "MFB-CBRNN: A hybrid network for MI detection using 12-lead ECGs," *IEEE Journal of Biomedical and Health Informatics*, Vol. 24, 2020, pp. 503-514.
26. W. Liu, M. Zhang, Y. Zhang, *et al.*, "Real-time multilead convolutional neural network for myocardial infarction detection," *IEEE Journal of Biomedical and Health Informatics*, Vol. 22, 2018, pp. 1434-1444.
27. C.-C. Lin and C.-M. Yang, "Heartbeat classification using normalized RR intervals and morphological features," *Mathematical Problems in Engineering*, Vol. 2014, 2014, pp. 1-11.
28. "Electrical conduction system of the heart," https://en.wikipedia.org/wiki/Electrical_conduction_system_of_the_heart.
29. E. Lee, E.-K. Choi, K.-D. Han, *et al.*, "Mortality and causes of death in patients with atrial fibrillation: A nationwide population-based study," *PLOS One*, Vol. 13, 2018, pp. 1-14.
30. C. A. Morillo, A. Banerjee, P. Perel, *et al.*, "Atrial fibrillation: the current epidemic," *Journal of Geriatric Cardiology*, Vol. 14, 2017, pp. 195-203.
31. E. J. Benjamin, P. A. Wolf, R. B. D'Agostino, *et al.*, "Impact of atrial fibrillation on the risk of death," *Circulation*, Vol. 98, 1998, pp. 946-952.
32. S. Stewart, C. L. Hart, D. J. Hole, and J. J. McMurray, "A population-based study of the long-term risks associated with atrial fibrillation: 20-year follow-up of the Renfrew/Paisley study," *The American Journal of Medicine*, Vol. 113, 2002, pp. 359-364.
33. "Holter monitor," <https://www.heart.org/en/health-topics/heart-attack/diagnosing-a-heart-attack/holter-monitor>, 2020.
34. "Comparison and review of portable, handheld, 1-lead/channel ECG / EKG recorders," <https://www.ndsu.edu/pubweb/~grier/Comparison-handheld-ECG-EKG.html>, 2020.
35. Y. Bai, J. Xie, C. Liu, Y. Tao, B. Zeng, and C. Li, "Regression modeling for enterprise electricity consumption: A comparison of recurrent neural network and its variants," *International Journal of Electrical Power & Energy Systems*, Vol. 126, 2021, No. 1106612.
36. M. E. Basiri, S. Nemati, M. Abdar, *et al.*, "ABCDM: An attention-based bidirectional CNN-RNN deep model for sentiment analysis," *Future Generation Computer Systems*, Vol. 115, 2021, pp. 279-294.
37. M. Abdel-Nasser and K. Mahmoud, "Accurate photovoltaic power forecasting models using deep LSTM-RNN," *Neural Comput & Applic*, Vol. 31, 2019, pp. 2727-2740.
38. "Tf.keras.losses.CategoricalCrossentropy, Tensorflow Core v2.6.0," *TensorFlow*, https://www.tensorflow.org/api_docs/python/tf/keras/losses/CategoricalCrossentropy, 2021.
39. "Tf.keras.metrics.accuracy, Tensorflow Core v2.6.0," *TensorFlow*, https://www.tensorflow.org/api_docs/python/tf/keras/metrics/Accuracy, 2021.

40. Brownlee, "Use early stopping to halt the training of neural networks at the right time," *Machine Learning Mastery*, <https://machinelearningmastery.com/how-to-stop-training-deep-neural-networks-at-the-right-time-using-early-stopping/>, 2021.
41. Brownlee, "A gentle introduction to mini-batch gradient descent and how to configure batch size," *Machine Learning Mastery*, <https://machinelearningmastery.com/gentle-introduction-mini-batch-gradient-descent-configure-batch-size/>, 2021.
42. J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE Transactions on Biomedical Engineering*, Vol. BME-32, 1985, pp. 230-236.
43. H. Khamis, R. Weiss, Y. Xie, C. W. Chang, N. H. Lovell, and S. J. Redmond, "QRS detection algorithm for telehealth electrocardiogram recordings," *IEEE Transactions on Biomedical Engineering*, Vol. 63, 2016, pp. 1377-1388.
44. M. Martinez-Selles, *et al.*, "Should we anticoagulate patients at high risk of atrial fibrillation?" *Revista Espanola de Cardiologia* (English edition), Vol. 69, 2016, pp. 374-376.



Agung Tri Arta Hariyanto (盧成江) received the B.Eng. degree in Informatics Engineering from Petra Christian University. He received the MS degree from the Department of Computer Science, National Chiao Tung University. His research interests include fog computing, cloud computing, and deep learning for medical diagnosis.



Kuochen Wang (王國禎) received the BS degree in Control Engineering from the National Chiao Tung University, Taiwan, in 1978, and the M.S. and Ph.D. degrees in Electrical Engineering from the University of Arizona in 1986 and 1991, respectively. He is currently a Professor of Center for Fundamental Science and was the Vice President for Library and Information Services from August 2019 to August 2021, Kaohsiung Medical University. He was a Professor and the Chair of the Department of Computer Science, National Chiao Tung University. He was a Director in the Institute of Computer Science and Engineering, Institute of Network Engineering, National Chiao Tung University from August 2009 to July 2011. He was an Acting/Deputy Director of the Computer and Network Center at this university from June 2007 to July 2009. He was a Visiting Scholar in the Department of Electrical Engineering, University of Washington from July 2001 to February 2002. From 1980 to 1984, he was a Senior Engineer at the Directorate General of Telecommunications in Taiwan. He served in the army as a second lieutenant communication platoon leader from 1978 to 1980. His research interests include AI and machine learning, IoT and big data analytics, and cloud fog computing and SDN.



Wei-Chung Tsai (蔡維中) received his M.D. and MSc degrees from Kaohsiung Medicine University during 1996-2010. He is currently an Assistant Professor, Department of Internal Medicine, College of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan. His research interests include arrhythmia, electrophysiology study and radiofrequency ablation.



Chi-Tsung Chen (陳季聰) received the MS and Ph.D. degrees in the Department of Aeronautics and Astronautics Engineering from National Cheng Kung University, Taiwan. His research interests include AI, machine learning, and nonlinear dynamics. He is currently a Research Fellow of the Office of Library and Information Service in the Kaohsiung Medicine University.