BeatClass: A Sustainable ECG Classification System in IoT-based eHealth

Le Sun*, Yilin Wang, Zhiguo Qu, and Neal N. Xiong*

***Abstract*—With the rapid development of Internet of Things (IoT), it becomes convenient to use mobile devices to remotely monitor the Physiological signals (e.g., Arrhythmia diseases) of patients with chronic diseases (e.g., cardiovascular diseases (CVDs)). High classification accuracy of inter-patient ECGs is extremely important for diagnosing Arrhythmia. The Supraventricular ectopic beat (S) is especially difficult to be classified. It is often misclassified as Normal (N) or Ventricular ectopic beat (V). Class imbalance is another common and important problem in electronic health (eHealth), as abnormal samples (i.e., samples of specific diseases) are usually far less than normal samples. To solve these problems, we propose a sustainable deep learning-based heartbeat classification system, called BeatClass. It contains three main components: two stacked bidirectional Long Short-Term Memory Networks (Bi-LSTMs), called Rist and Morst, and a Generative Adversarial Network (GAN), called MorphGAN. Rist first classifies the heartbeats into five common Arrhythmia classes. The heartbeats classified as S and V by Rist are further classified by Morst to improve the classification accuracy. MorphGAN is used to augment the morphological and contextual knowledge of heartbeats in infrequent classes. In the experiment, BeatClass is compared with several state-of-the-art works for inter-patient arrhythmia classification. The F1-scores of classifying N, S and V heartbeats are 0.6%, 16.0% and 1.8% higher than the best baseline method. The experiment result demonstrates that taking multiple classification models to improve classification results step-by-step may significantly improve the classification performance. We also evaluate the classification sustainability of BeatClass. Based on different physical signal datasets, a trained BeatClass can be updated to classify heartbeats with different sampling rates. At last, an engineering application indicates that BeatClass can promote the sustainable development of IoT-based eHealth.**

***Index Terms*—IoT-based eHealth, ECG Classification, Sustainable System.**

I. INTRODUCTION

GOOD health and wellbeing is a current focal point for addressing the sustainable development goals [1]. The fast development of Internet of Things (IoTs) technology has brought new opportunities and challenges for electronic Health (eHealth) [2]. The development of wireless sensor network [3] and data mining technology [4] has greatly promoted the progress of eHealth. Intelligently processing the massive remotely monitored healthcare data is getting an increasing amount of academic attention [5]. Telemedicine for online disease diagnosis can assist doctors to make diagnostic decisions in advance, e.g., combining the advantages of edge and cloud computing to collect and process ECG data [6]. Patients in remote areas can also enjoy medical treatment at home, which saves fuel for travel, and lessens the cost of medical treatment [7] [8]. Therefore, developing an effective Telemedicine system for online disease diagnosis can dramatically promotes the sustainable development of IoT [9].

ElectroCardioGram (ECG) is a type of IoT-based eHealth data that is easy to be accessed and collected by using wearable devices or mobile terminals [10]. It is an effective means for diagnosing the cardiovascular diseases [11]. Arrhythmia prevents the heart from pumping enough blood into vital organs [12]. Early detection and timely treatment are essential. The Association for Advancement of Medical Instrumentation (AAMI) [13] classifies the heartbeats of the arrhythmia patients into five classes: Normal beat (N), Supraventricular ectopic beat (S), Ventricular ectopic beat (V), Fusion beat (F), and Unclassifiable beat (Q). Table I shows the heartbeat classes contained in each class. Accurately classifying arrhythmia heartbeats is a hot and difficult research topic in eHealth [12].

Several researches explore the intelligent classification of

TABLE I: Five Classes of Heartbeats Defined by AAMI

Class	Annotation	Heartbeat
N	N	Normal beat
	L	Left bundle branch block beat
	R	Right bundle branch block beat
S	A	Atrial premature beat
	J	Nodal (junctional) premature beat
	S	Supraventricular premature beat
	a	Aberrated atrial premature beat
	e	Atrial escape beat
	j	Nodal (junctional) escape beat
V	E	Ventricular escape beat
	V	Premature ventricular contraction
F	F	Fusion of ventricular and normal beat
Q	P	Paced beat
	Q	Unclassifiable beat
	f	Fusion of paced and normal beat

Manuscript received April 19, 2005; revised August 26, 2015.

Yilin Wang, Le Sun and Zhiguo Qu are with the Engineering Research Center of Digital Forensics, Ministry of Education, Nanjing University of Information Science and Technology; and the Department of Jiangsu Collaborative Innovation Center of Atmospheric Environment and Equipment Technology (CICAEET), Nanjing University of Information Science and Technology, Nanjing, 210044, China. Email: 20191220029@nuist.edu.cn, LeSun1@nuist.edu.cn and qzghhh@126.com.

Neal N. Xiong is with the Department of Computer Science and Mathematics, Sul Ross State University, Alpine, TX 79830, USA. Email: xiongnaixue@gmail.com.

*: Both Le Sun and Neal N. Xiong are the corresponding authors.

the five-class heartbeats [14]–[16]. From a technical perspective, there are two types of classification methods: traditional machine learning methods [14], [17] and deep learning methods [15], [18], [19]. Given enough data, the classification performance of traditional methods is usually worse than that of the deep learning methods [20]. From the perspective of experimental subjects, two types of researches were conducted to classify inter-patient [21]–[23] and intra-patient [15], [24] heartbeats respectively. *Inter-patient* means a classification model can classify heartbeats of different patents. *Intra-patient* means a model only classifies heartbeats of a patient. Inter-patient classification is more difficult than intra-patient classification [19]. Our work focuses on the inter-patient classification of the five-class heartbeats based on deep learning methods.

Data imbalance is a common problem in eHealth, as abnormal samples (e.g., disease samples) are normally far less than normal samples. [25]. It is a negative factor influencing the classification performance [26]. There are usually many normal (N) heartbeats but fewer abnormal (e.g., V, S, F and Q) heartbeats in an ECG dataset [24]. Acharya et al. [15] uses a statistical method to augment S, V, F and Q heartbeats. He et al. [12] designed an S heartbeat augmentation algorithm based on medical knowledge. Their data augmentation algorithm strongly relies on expert knowledge and can only augment S heartbeats. A more general method is to use the Generative Adversarial Network (GAN) for heartbeat augmentation [27]. In this work, we design a GAN-based data augmentation method, called MorphGAN, to augment the Contbeat and RRs of infrequent-class heartbeats.

Based on the classification results of previous works [28] [29] [30], it is difficult to classify S heartbeats. They are often misclassified as N or V heartbeats. For example, for S, V and N classification, the F1-scores of work [28] are 45.5%, 79.5% and 95.3% respectively. The F1-scores of [30] are 60.8%, 81.5% and 90.0% respectively. The F1-scores of [31] are 70.1%, 85.5% and 88.9% respectively. To solve this problem, we take advantage of the clinical manifestations of heartbeats [12] in classification: (1) the Morphologies (Mors) of N and S heartbeats are similar to each other; (2) the RR interval (RR) of an S heartbeat is normally shorter than that of an N heartbeat; and (3) the RRs of S and V heartbeats are similar to each other. An RR is the length between the R waves of the current heartbeat and its previous adjacent heartbeat (see Fig. 1). The RRs and Mors are two key features for classifying heartbeats [12]. Classifying F and Q is not a research focus because their classification are meaningless [12]. We will not discuss F and Q in-depth.

Based on the clinical manifestations, our previous work [12] develops a two-step classification system to classify N, S and V heartbeats step-by-step. It demonstrates that using more than one models to improve the classification results can achieve good performance. Inspired by [12], we develop a new two-step heartbeat classification system, called **BeatClass**. It contains three main components: a stacked Bi-LSTM (called Rist) to classify the ECG heartbeats into four classes (N, S&V, F and Q) based on the RR; a stacked Bi-LSTM (called Morst) to further improve the classification of S and V based on the Mor and the **Cont**extual information of heart**beat**s (Contbeat); and

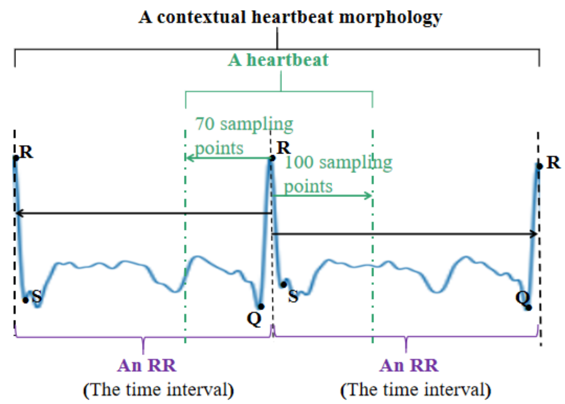


Fig. 1: Heartbeat, RR and Contbeat.

a GAN-based data augmentation method (called MorphGAN) to augment the Contbeat of the heartbeats in infrequent classes. The Contbeat of a heartbeat contains two adjacent heartbeats (i.e., just before and after) of the heartbeat (see Fig. 1). BeatClass extends [12] from four aspects: (1) designs two new heartbeat classification models; (2) proposes a new feature Contbeat to capture the contextual information of a heartbeat; (3) proposes a GAN-based data augmentation method to solve the data imbalance problem; and (4) evaluates the classification performance of the proposed models based on more datasets. Experiment results show that BeatClass performs much better than [12]. Our distinctive contribution is as follows:

- Propose a new two-step deep learning-based classification system, BeatClass, to classify inter-patient heartbeats. It demonstrates that using multiple classification models to improve the classification results step-by-step may dramatically improve the classification performance.
- Design a new feature Contbeat. It is useful for classifying V and S heartbeats. By considering such contextual information, the heartbeat classification performance is further improved, especially the classification of S and V heartbeats.
- Propose a new GAN (MorphGAN) to augment the features (e.g., Contbeat and RR) of the heartbeats in infrequent classes.
- Conduct comprehensive experiments. BeatClass is compared with several state-of-the-art works for inter-patient arrhythmia classification. The overall accuracy of BeatClass is 98.7%, which is 27.4% and 0.03% higher than the worst and best baseline methods respectively. The F1-scores of classifying N, S and V heartbeats are 99.5%, 94.7% and 97% respectively, which are 0.6%, 10.7% and 1.8% higher than the best baseline methods.
- Present the sustainability of BeatClass from two aspects: (1) use four different datasets to show the classification capability of BeatClass is sustainable. Based on new ECG dataset, a trained BeatClass can be updated to classify the dataset with different sampling rates. This sustainable capability makes BeatClass more robust in the dynamic IoT-based Telemedicine. To our knowledge, we are the first to validate the sustainability of heartbeat classifica-

TABLE II: Comparison of State-of-the-art Deep Learning Methods for Heartbeat Classification

Work	Step-by-step Improvement of Classification Results	Contextual Information	Sustainability	F1 Score of Classifying S Heartbeats (%)
Chazal et al. [14]	No	The previous RR interval, post RR interval, average RR interval and the local RR interval	No	51.5
Niu et al. [23]	No	The previous RR interval	No	76.6
Li et al. [21]	No	Not considered	No	79.9
Saadatnejad et al. [32]	No	Not considered	No	78.8
Li et al. [33]	No	The ratio of the current pre-RR interval to the average of all pre-RR intervals	No	78.4
Mohamed et al. [34]	No	Not considered	No	80.1
BeatClass	Yes	Consider an RR interval sequence contains 50 RRs and the contextual relationship between adjacent RR intervals	Yes	94.7

tion based on four ECG datasets; and (2) use a practical application of BeatClass to show how it improves the sustainable development of IoT-based Telemedicine.

The rest of this paper is organized as follows: Section II introduces the related work; Section III presents the formal definition and the detailed structure of BeatClass; Section IV shows the performance analysis and an engineering application in IoT-based eHealth; and Section V concludes this paper and gives some future research directions.

II. Related Work

Automatic classification of arrhythmia based on ECG has been a hot research topic in the field of health informatics. There are researches solving either intra-patient [15] [24] or inter-patient [14] [17] [12] [32] heartbeat classification. Our discuss will focus on inter-patient classification.

A. Arrhythmia Heartbeat Classification: Multi-task Classification

1) Traditional Machine Learning Classification Methods: Ye et al. [17] proposed a new method based on support vector machine (SVM) for heartbeat classification using morphological and dynamic characteristics. Chen et al. [35] proposed a heartbeat classification based on SVM with the combination of projected and dynamic features. Zhang et al. [36] developed a disease feature selection method consisting of a one-versus-one feature sequencing stage and a feature search stage. Work [36] does not solve the problem of data imbalance and has no verification of universality. Llamedo and Martínez [37] proposed a simple ECG classifier based on ECG feature model selection, with the emphasis on improving its generalization ability. Mondéjar-Guerra et al. [38] proposed an automatic classification method based on multi-SVMs. This method relies on the time interval between heartbeats and their morphology. Shi et al. [39] developed a hierarchical classification method based on weighted extreme gradient boosting. Lee et al. [40] proposed a novel algorithm for hybrid neuro-fuzzy logic system based on local transformation pattern.

2) Deep Learning Classification Methods: We compare the deep learning-based heartbeat classification from four aspects (see Table II): (1) using more than one model to improve the classification results step-by-step; (2) consider contextual information; (3) evaluate the sustainability of a classification model on multiple datasets; and (4) average classification accuracy.

Most previous works just use one comprehensive model to classify heartbeats. Chazal et al. [14] selected the optimal linear discriminate classifier configuration based on the feature set of ECG morphology, heartbeat interval and RR. Niu et al. [23] proposed a heartbeat classification method based on Multi-Perspective Convolutional Neural Network (MPCNN). Li et al. [21] proposed a method based on General Convolutional Neural Network (GCNN) to classify heartbeats. Saadatnejad et al. [32] proposed an LSTM-based heartbeat classification system. Li et al. [33] proposed a high performance automatic heartbeat classification method based on CNN. Mohamed et al. [34] proposed a classification strategy based on the deep neural network. The learning stage of this method is based on a powerful feature extraction protocol to improve the classification accuracy. Different from the above works, BeatClass uses two classification models to improve the classification results step-by-step. The average classification accuracy of BeatClass is higher than other works (see the last column in Table II).

ECG is a type of time series data. The contextual information of a heartbeat usually contains its neighbor relation and the temporal information. Thus, considering contextual information is also important in heartbeat classification. Chazal et al. [14] used the contextual information: the previous RR interval, post RR interval, average RR interval and the local RR interval of a heartbeat in classification. Niu et al. [23] and Li et al. [33] considered the previous RR intervals of heartbeats. Our system BeatClass uses an RR interval sequence contains 50 RRs as a feature for classification. We also design a new feature Contbeat to capture the contextual and temporal relation between a heartbeat and its neighbors.

Sustainability indicates the capability of a classification model classifying ECG datasets with different sampling rates. Most works summarized in Table II do not evaluate the sustainability of their models. Few of them test their models on classifying multiple datasets. In our work, we train BeatClass on MIT-BIH-AR dataset [14]. The trained model is then updated to classify heartbeats in three other datasets INCART [41], MIT-BIH-SUP [42] and MIT-BIH-L [43]. Experiment results show that BeatClass performs excellent on the different datasets, which demonstrate the good sustainability of BeatClass.

The last column of Table II shows the F1 score of classifying S heartbeats based on MIT-BIH-AR dataset [14]. BeatClass significantly improves the classification performance for S heartbeats compared with the State-of-the-Art methods.

B. Data Augmentation Methods

1) GAN-Based Augmentation Methods: Shaker et al. [44] proposed an end-to-end deep learning method and a two-stage hierarchical learning method based on deep CNNs. These methods can combine feature extraction, feature reduction and classification into a single learning method. GAN based on fully connected layers is used to restore the balance of the dataset. Wang et al. [45] designed a classification model based on stacked residual network combined with LSTM. The data augmentation method uses auxiliary classifier generative adversarial network (ACGAN) based on CNN. Zhou et al. [46] proposed a GAN-based high performance automatic arrhythmia classification system. The generator of GAN is used to generate coupling matrices for data augmentation. After training, the discriminator is extracted as an arrhythmia classifier. Golany et al. [47] proposed an ordinary differential equation (ODE) system representing cardiac dynamics. And the ODE system is incorporated into the optimization process of GAN to generate ECG data. Golany and Radinsky [27] proposed a semi-supervised approach for patient-specific ECG classification. A GAN-based generation model is developed to learn and synthesize patient-specific ECG signals to improve the performance of patient-specific classifiers.

2) Other Data Augmentation Methods: Acharya et al. [15] designed a nine-layer CNN for classification and a data augmentation method based on Z-score normalization. Rajesh and Dhuli [24] proposed a nonstationary nonlinear decomposition technique, called improved Fully Integrated Empirical Mode decomposition (ICEEMD), to extract features. Features are then fed into the AdaBoost integrated classifier for heartbeat classification. In addition, three data augmentation techniques: Re-sampling, Synthetic Minority Oversampling Technique (SMOTE), and Distribution based data sampling are used. Sellami and Hwang [48] proposed a new deep CNN to accurately classify the heartbeat. In order to overcome the imbalance between classes, a batch weighted loss function is proposed to better quantify the loss.

C. Abnormal Heartbeat Detection: Binary Classification

Atrial Fibrillation (AF) is a common chronic arrhythmia. AF can cause heart failure, blood clots and other diseases. Early detection and timely treatment of AF can effectively avoid disability or death. Part of the research is devoted to the detection of AF. These studies can screen out AF. Zhou et al. [49] proposed a GAN-based algorithm BeatGAN for generating and detecting abnormal heartbeats. Shashikumar et al. [50] proposed a deep learning system based on attention. Firstly, the deep convolutional neural network is used to extract the image features. And the image features are then extracted and presented to the bidirectional RNN with attention layer for AF detection.

There are also some studies that classify ECG records. The classification categories are Normal Sinus Rhythm (NSR), AF, other types of arrhythmia except AF (other), noise signal (noise). Cao et al. [51] proposed an improved multi-scale decomposition enhanced residual CNN. This model does not need complex preprocessing methods and can be generalized to classify other ECG datasets. Dang et al. [52] proposed a network model based on CNN and Bi-LSTM to automatically detect AF heartbeat. The model is composed of four convolution layers, two Bi-LSTM layers and two fully connection layers.

Works [21]–[24], [32], [38]–[40] fail to verify the generalizability of their classification methods in other data sets. Our work is different to the above works, because our work simultaneously solves the following problems: (1) classify the five classes of Arrhythmias heartbeats based on the new AAMI standard [13]; (2) verify the generalizability of the proposed classification method based on different ECG datasets that have different data distributions; and (3) augment the heartbeats in small classes by designing a new GAN model. Experiment results show that our classification model BeatClass performs much better than the models in [12], [14], [15], [17], [21]–[24], [32], [35]–[40] (see Section IV).

III. Our Proposed BeatClass System

A. Concepts, Notations and BeatClass Overview

The concepts of a heartbeat, an RR-interval and a Contbeat are shown in Fig. 1. They are defined based on the ECG signals in 360Hz sampling rate. A **heartbeat** $hb_k = \{x_{k1}, x_{k2}, \cdots, x_{kH}\}$ contains 70 and 100 sampling points before and after an R-wave respectively [12]. The length of hb_k is represented by $|hb_k|$ = H = 170. x_{kh} (h $\in$ [1,H]) is a signal value at the h_{th} time tick of the k_{th} heartbeat. This period can cover the P-wave, QRS-wave and T-wave of a heartbeat [12]. An **RR-interval** r_k (abbreviated as RR) is a time interval between the R-wave of the k_{th} heartbeat and the R-wave of the $(k-1)_{th}$ heartbeat [20]. A **Contbeat** $X_k = \{x_k^1, x_k^2, \cdots, x_k^T\}$ represents the contextual morphology of the k_{th} heartbeat. A Contbeat contains 180 and 220 sampling points before and after an R-wave respectively. x_k^j (j $\in$ [1,T], T=400) is a signal value at the j_{th} time tick of the k_{th} Contbeat. The length of X_k is represented by $|X_k|$.

The **RR sequence** of the k_{th} heartbeat (hb_k) is defined as $s_k = \{r_{k-L+1}, \cdots, r_{k-1}, r_k, r_{k+1}, \cdots, r_{k+L}\}$. r_m $(m \in [k-L+1, \cdots, k+L])$ is the RR of the m_{th} heartbeat. The normalized RR sequence is represented by $S_k = \{R_{k-L+1}, \cdots, R_{k-1}, R_k, R_{k+1}, \cdots, R_{k+L}\}$, where, $R_i = r_i/avepre$ $(i \in [k-L+1, \cdots, k])$; $R_j = r_j/avepos$ $(j \in [k+1, \cdots, k+L])$; L is a hyper-parameter (see Table III); $avepre$ and $avepos$ are the means of the first L RRs and the last L RRs in s_k respectively.

TABLE III: Meaning of the Main Notations

Notation	Meaning
$hb_k = \{x_{k1}, x_{k2}, \cdots, x_{kH}\}$	The k_{th} heartbeat; H is the length of a heartbeat
r	An RR value of the RR sequence before normalization
R	An RR value of the RR sequence after normalization
$s_k = \{r_{k-L+1}, \cdots, r_{k-1}, r_k, r_{k+1}, \cdots, r_{k+L}\}$	s_k is the RR sequence of hb_k; r_m ($m \in [k-L+1, \cdots, k+L]$) is the RR of the m^{th} heartbeat
L	The range of the heartbeats
$S_k = \{R_{k-L+1}, \cdots, R_{k-1}, R_k, R_{k+1}, \cdots, R_{k+L}\}$	The RR sequence after normalization
$\mathbf{S} = [S_1, S_2, \cdots, S_N] \in \mathbb{R}^{N\times 1\times C}$	$\mathbf{S}$ is a collection of RR sequences; N is the number of RR sequences in $\mathbf{S}$; C is the number of RRs in each RR series
$X_k = \{x_k^1, x_k^2, \cdots, x_k^T\}$	X_i is the i^{th} Contbeat; x_k^j ($j \in [1, \cdots, T]$) is a ECG signal value at the j^{th} time tick of the X_i; T is the number of segmented sampling points
$\mathbf{X} = [X_1, X_2, \cdots, X_n] \in \mathbb{R}^{n\times 1\times T}$	$\mathbf{X}$ is the collection of Contbeats; n is the number of Contbeats in $\mathbf{X}$
$Z_k = \{z_k^1, z_k^2, \cdots, z_k^t\}$	A random noise signal satisfying the 0-1 Gaussian distribution; t is the length of Z_k
$\lvert\cdot\rvert$	Length of a signal. For example, $\lvert Z_k\rvert$ means the number of points in the noise signal Z_k

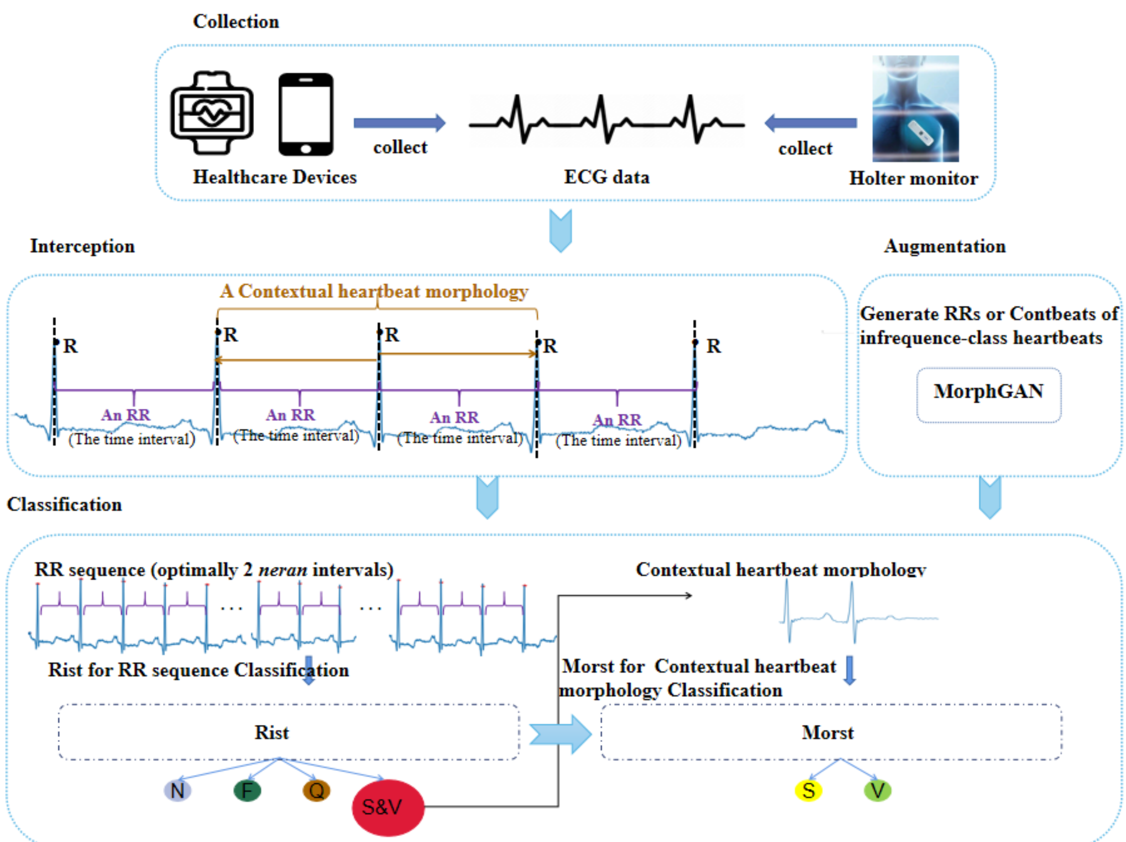


Fig. 2: BeatClass: a sustainable ECG classification system.

The meanings of main notations used in BeatClass are summarized in Table III.

BeatClass is a sustainable arrhythmia detection system (see Fig. 2 and Alg. 1). It consists of two main steps. In the first step, a Stacked Bi-LSTM (called Rist) classifies the real heartbeats into four classes: N, S&V, F and Q, based on RR sequences (lines 4-6 in Alg. 1). The heartbeats classified into S&V class will be further classified by the second step (lines 7-10 in Alg. 1). In the second step, a Stacked Bi-LSTM (called Morst) distinguishes S from V based on the Contbeats. In addition, a GAN model, called MorphGAN, is used to augment the Contbeats or RRs of the infrequent-class heartbeats.

Algorithm 1 BeatClass(hb_k, L)

1: Input: a heartbeat hb_k, L;
2: Output: c=the class of hb_k;
3: Algorithm start:
4: Find RR sequence of hb_k: $s_k = \{r_{k-L+1}, \cdots, r_{k-1}, r_k, r_{k+1}, \cdots, r_{k+L}\}$;
5: Normalize s_k: $S_k = \{R_{k-L+1}, \cdots, R_{k-1}, R_k, R_{k+1}, \cdots, R_{k+L}\}$;
6: u =($P_N, P_{S\&V}, P_F, P_Q$)=Rist(S_k);
7: **if** $P_{S\&V}$=max(u) **then**
8: Find Contbeat of hb_k: $X_k = \{x_k^1, x_k^2, \cdots, x_k^T\}$;
9: u=(P_S,P_V) = Morst(X_k);
10: **end if**
11: return c = $\underset{c\in\{N,S,V,F,Q\}}{\arg\max}$ (u);

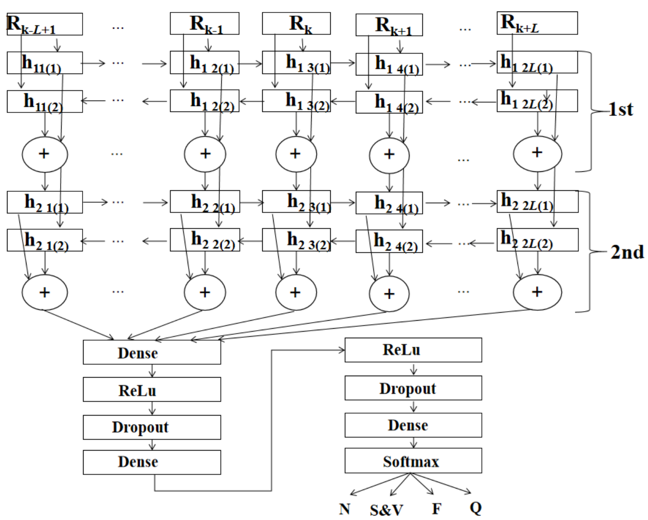


Fig. 3: Structure of Rist.

B. Rist in BeatClass

The structure of Rist is shown in Fig. 3. Rist is trained based on **S** = [S_1, S_2, $\cdots$, S_N], where N is the number of RR sequences in the training set, and S_k is the normalized RR sequence of the classified heartbeat hb_k, k $\in$ [1,N]. The main structure of Rist is a Bi-LSTM. Each LSTM layer has $2L$ cells. The input of the forward and backward LSTM layers are $S_k = \{R_{k-L+1}, \cdots, R_{k-1}, R_k, R_{k+1}, \cdots, R_{k+L}\}$ and its reversed signal $S'_k = \{R_{k+L}, \cdots, R_{k+1}, R_k, R_{k-1}, \cdots, R_{k-L+1}\}$ respectively. The number of cells in Rist is shown in Table V. The parameter settings of Rist are in Table IV. Rist combines the stacked Bi-LSTM with a fully connected layer. After the fully connected layer, Rist adds several dropout layers with dropout rate of 0.3 to prevent overfitting, and uses ReLu as the activation function. It outputs the probability of hb_k belonging to N, S&V, F, and Q classes (line 6 in Alg. 1).

Eqs. (1) to (6) are the mathematical representation of the Bi-LSTM in Rist. Each equation has two sub-equations. Sub-equations with the lower index of *ft* and *bt* represent the processing processes of the forward and backward layers respectively. Eq. (1) determines the information to be forgotten. Eq. (2) determines the information to be updated in the forward and backward layers. Eq. (3) determines the alternate content C_{ft} and C_{bt} to be updated. $\tilde{C}_{ft}$ and $\tilde{C}_{bt}$ are optional for updating. Eq. (4) adds new information to the cell states C_{ft} and C_{bt} to replace the old cell states C_{ft-1} and C_{bt-1} respectively. The output gate selects the important information from the current state as the output of the cellular state. Eq. (5) determines o_{ft} and o_{bt}, i.e., the output information. Eq. (6) transforms o_{ft} and o_{bt} to the output h_{ft} and h_{bt} respectively. Finally, (7) combines the output of the forward layer and the backward layer in (6) to get the final output O_t.

TABLE IV: Parameter Settings of Rist and Morst

Parameter	Settings
Dropout rate	0.3
Training epoch	50 (Rist) & 100 (Morst)
Optimizer	Adam [53]
Loss function	The sparse categorical crossentropy
Learning rate	0.001 [54]

$$\begin{aligned} f_{ft} &= \sigma(W_{ff} \cdot [h_{ft-1}, R_t] + b_{ff}), \\ f_{bt} &= \sigma(W_{bf} \cdot [h_{bt-1}, R_t] + b_{bf}), \end{aligned} \tag{1}$$

$$\begin{aligned} i_{ft} &= \sigma(W_{fi} \cdot [h_{ft-1}, R_t] + b_{fi}), \\ i_{bt} &= \sigma(W_{bi} \cdot [h_{bt-1}, R_t] + b_{bi}), \end{aligned} \tag{2}$$

$$\begin{aligned} \tilde{C}_{ft} &= tanh(W_{fc} \cdot [h_{ft-1}, R_t] + b_{fc}), \\ \tilde{C}_{bt} &= tanh(W_{bc} \cdot [h_{bt-1}, R_t] + b_{bc}), \end{aligned} \tag{3}$$

$$\begin{aligned} C_{ft} &= f_{ft} * C_{ft-1} + i_{ft} * \tilde{C}_{ft}, \\ C_{bt} &= f_{bt} * C_{bt-1} + i_{bt} * \tilde{C}_{bt}, \end{aligned} \tag{4}$$

$$\begin{aligned} o_{ft} &= \sigma(W_{fo} \cdot [h_{ft-1}, R_t] + b_{fo}), \\ o_{bt} &= \sigma(W_{bo} \cdot [h_{bt-1}, R_t] + b_{bo}), \end{aligned} \tag{5}$$

$$\begin{aligned} h_{ft} &= o_{ft} * tanh(C_{ft}), \\ h_{bt} &= o_{bt} * tanh(C_{bt}), \end{aligned} \tag{6}$$

$$O_t = \sigma(W_O[h_{ft}, h_{bt}] + b_O). \tag{7}$$

where, σ and $tanh$ are activation functions Sigmoid and Tanh respectively. R_t is the t^{th} ($t \in [k - L + 1, \cdot \cdot \cdot, k + L]$) RR in an RR sequence. In the forward layer, W_{ff}, W_{fi}, W_{fc} and W_{fo} are the weights between neurons; b_{ff}, b_{fi}, b_{fc} and b_{fo} are the biases in the forgotten gate, input gate, update process and output gate, respectively. In the backward layer, W_{bf}, W_{bi}, W_{bc} and W_{bo} are the weights between neurons; b_{bf}, b_{bi}, b_{bc} and b_{bo} are the biases in the forgotten gate, input gate, update process and output gate, respectively. W_O and b_O are the weight and bias of the final output.

C. Morst in BeatClass

Morst is used to further classify S and V heartbeats (line 9 in Alg. 1). The structure of Morst is shown in Fig. 4. Its structure is similar with the structure of Rist, but the input and the number of cells in each layer (Table VI) are different. The input of Morst is a Contbeat (line 8 in Alg. 1). We denote all

TABLE V: Number of Cells in Rist

Layer	The number of cells	Output shape
1st LSTM	50	(404, 1, 100)
2nd LSTM	50	(604, 100)
1st Dense	64	(101, 64)
Dropout layer	-	(101, 64)
2nd Dense	32	(65, 32)
Dropout layer	-	(65, 32)
3rd Dense	4	(33, 4)

TABLE VI: Number of Cells in Morst

Layer	The number of cells	Output shape
1st LSTM	400	(3204, 1, 800)
2nd LSTM	400	(4804, 800)
1st Dense	64	(801, 64)
Dropout Layer	-	(801, 64)
2nd Dense	32	(65, 32)
Dropout Layer	-	(65, 64)
3rd Dense	2	(33, 2)

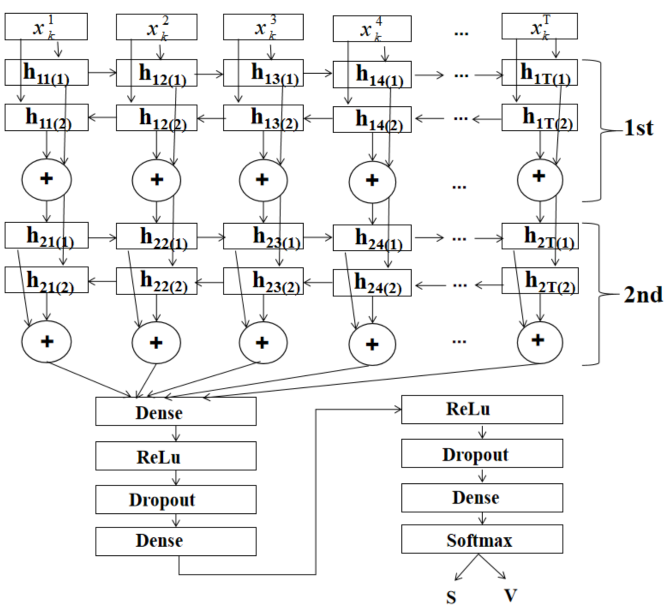


Fig. 4: Structure of Morst.

collected Contbeats as $\mathbf{X}$ = $[X_1,\ X_2,\ \cdots,\ X_n] \in \mathbb{R}^{n\times 1\times T}$, where n is the number of Contbeats. The parameter settings of Morst are shown in Table IV.

D. MorphGAN for Augmenting Classes with Small Sample Sizes

MorphGAN is used to augment the infrequent-class Contbeats to solve the class imbalance problem in classifier training.

The structure of MorphGAN is shown in Fig. 5. The generator (G) generates realistic ECG data, and the discriminator (D) judges whether the input ECG data is real. The parameter settings of D and G are shown in Table VII and

TABLE VII: Structure of Discriminator in MorphGAN

Layer	Type	Kernel size	Stride
0	Input heartbeat	-	-
1	Conv1D	(32,6)	3
2	Batch normalization	-	-
3	LeakyReLU	-	-
4	MaxPooling1D	2	-
5	Flatten	-	-
6	Dense (Sigmoid)	Unit(1)	-

TABLE VIII: Structure of Generator in MorphGAN

Layer	Type	Kernel size	Stride
0	Input noise	-	-
1	Dense (ReLu)	Unit(128)	-
2	Reshape	-	-
3	UpSampling (2)	-	-
4	Conv1D	(128,6)	1
5	Batch normalization	-	-
6	ReLU	-	-
7	UpSampling (2)	-	-
8	Conv1D	(64,6)	1
9	Batch normalization	-	-
10	ReLU	-	-
11	UpSampling (2)	-	-
12	Conv1D	(32,6)	1
13	Batch normalization	-	-
14	ReLU	-	-
15	Conv1D	(1,6)	1
16	Sigmoid	-	-

Table VIII respectively. D has 32 six-dimension filters in Conv1D. MaxPooling layer has a two-dimension filter. The final Dense layer has a neuron. D contains a full connection layer, a convolution layer, a maximal pooling layer, and a flatten layer. The first Dense layer of G has 128 neurons. The number of six-dimension filters used in the four convolutional layers are 128, 64, 32, and 1, respectively. G has three repeated blocks. Each block has an Upsampling layer, a Convolutional layer, a Batch normalization layer and a ReLu layer. The momentum value of Batch normalization layers in D and G are 0.8. The alpha value of the LeakyReLu layer in D is 0.2. The number of neurons in the first fully connection layer is 6400. Through the reshape layer, the data size is changed to 50 $\times$ 128. After three rounds of upsampling and a convolution, the data size is changed to 400 $\times$ 1.

Table IX shows the parameter settings of MorphGAN. The input of D is a 400 $\times$ 1 vector, which can be Contbeats ($\mathbf{X}$) or RR sequences ($\mathbf{S}$) (see line 1 in Alg. 2). As $|S_k|$ = 50, we pad S_k with zeros. We use X_k ($|X_k|$=400) to represent the 400 $\times$ 1 input vector. The input of the generator G is noise signals $\mathbb{Z}$ = $\{Z_1,\ Z_2,\ \cdots,\ Z_{num}\}$, where Z_k = $\{z_k^1,\ z_k^2,\ \cdots,\ z_k^t\}$ (k $\in$ [1,num], t=100). Z_k is the k_{th} random noise satisfying the 0-1 Gaussian distribution. Eq. 8 shows the target function of

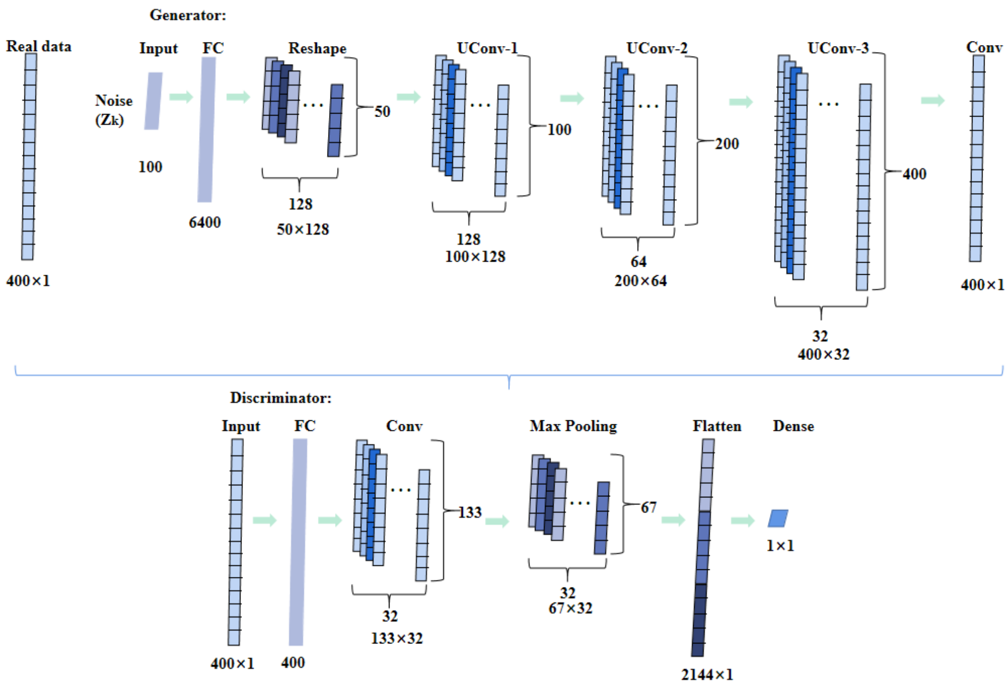


Fig. 5: Structure of MorphGAN.

Algorithm 2 MorphGAN Training

1: Input: $\mathbf{X} = [X_1, X_2, \cdots, X_n]$ (or $\mathbf{S} = [S_1, S_2, \cdots, S_n]$), a sequence of random noise $\mathbb{Z} = Z_1, Z_2, \cdots, Z_{num}$;
2: Output: trained MorphGAN;
3: **for** epochs **do**
4: **for** e steps **do**
5: Sample M noise samples $Z_1, Z_2, \cdots, Z_M$ from noise prior $P_g(\mathbb{Z})$;
6: Sample M examples $X_1, X_2, \cdots, X_M$ from data distribution $P_{data}(\mathbf{X})$;
7: Update the discriminator D by ascending its stochastic gradient: $\bigtriangledown_{\theta_d} \frac{1}{M} \sum_{k=1}^{M} [logD(X_k) + log(1 - D(G(Z_k)))]$.
8: **end for**
9: Sample M noise samples $Z_1, Z_2, \cdots, Z_M$ from noise prior $P_g(\mathbb{Z})$;
10: Update the generator G by descending its stochastic gradient: $\bigtriangledown_{\theta_g} \frac{1}{M} \sum_{k=1}^{M} [log(1 - D(G(Z_k)))]$.
11: **end for**
12: return trained MorphGAN.

MorphGAN.

$$\begin{aligned} \underset{G}{min}\, \underset{D}{max}\, V(D, G) = E_{\mathrm{X} \sim P_{data}(\mathrm{X})}[logD(\mathrm{X})] \\ + E_{\mathbb{Z} \sim P_{\mathbb{Z}}(\mathbb{Z})}[log(1 - D(G(\mathbb{Z})))]. \end{aligned} \tag{8}$$

where, $E_{\mathrm{X} \sim P_{data}(\mathrm{X})}$ represents the distribution of real data, and $E_{Z_k \sim P_{\mathbb{Z}}(\mathbb{Z})}$ represents the distribution of $\mathbb{Z}$.

Alg. 2 shows the training process of MorphGAN. The number of training epochs is 15000 (line 3). In each epoch, we update D e (e = 5) times (lines 4-8) and update G once (lines 9-10). M noise samples $Z_1, Z_2, \cdots, Z_M$ and M data samples $Z_1, Z_2, \cdots, Z_M$ are used to update D (lines 5-6). D is updated by ascending its stochastic gradient (lines 5 to 7). When updating G, a new set of noise samples $Z_1, Z_2, \cdots, Z_M$ from the G's distribution over noises $P_g(\mathbb{Z})$ are sampled (line 9). G is updated by descending its stochastic gradient (lines 9 to 10).

After the model training, a noise sequence $\mathbb{Z}' = Z'_1, Z'_2, \cdots, Z'_{num}$ (Z'_p ($p \in [1, 2, \cdots, num]$) is a noise signal) is used by MorphGAN to generate num RRs or Contbeats. The generated RRs or Contbeats will be mixed with real RRs or Contbeats to train Rist or Morst.

TABLE IX: Parameter Settings of MorphGAN

Parameter	Settings
The momentum value of Batch normalization layer	0.8
The alpha value of the LeakyReLu layer	0.2
Training epoch	15000
Optimizer	Adam
Loss function	The binary crossentropy
Learning rate	0.0001 (G) & 0.0002 (D)

IV. Performance Analysis

Experiments are conducted on five real ECG datasets to evaluate the performance of BeatClass. The experiments are run on a computer with a CPU of Intel(R) UHD Graphics 630 and 8.00 GB memory.

A. Datasets

Our experiments are based on five datasets:

TABLE X: Division of MIT-BIH-AR

Dataset	MIT-BIH-AR Recordings
DS1	101, 106, 108, 109, 112, 114, 115, 116, 118, 119, 122, 124, 201, 203, 205, 207, 208, 209, 215, 220, 223, 230
DS2	100, 103, 105, 111, 113, 117, 121, 123, 200, 202, 210, 212, 213, 214, 219, 221, 222, 228, 231, 232, 233, 234

- MIT-BIH Arrhythmia database (MIT-BIH-AR) contains 48 two-channel ECG recordings from Beth Israel hospital. The duration of each record is half an hour. The ECG signals are from 47 patients, which are digitized at a rate of 360 samples per second per channel [55]. Our experiment is based on lead MLII. The division follows the widely accepted patient-specific paradigm in [14], which divides MIT-BIH-AR into a training set (DS1) and a testing set (DS2). Table X shows the division.
- MIT-BIH Supraventricular Arrhythmia database (MIT-BIH-SUP) contains 78 half-hour ECG recordings to supplement the MIT-BIH-AR with examples of supraventricular arrhythmias. All the recordings are sampled at a frequency of 128 Hz [41]. Our experiment is based on the lead ECG1. Odd numbered records are the training set (called SDS1), and the rest are the testing set (called SDS2).
- St.-Petersburg Institute of Cardiological Technics 12-lead Arrhythmia database (INCART) consists of 75 records. Every record lasts 30 minutes and contains 12 standard leads. Each record has a sampling frequency of 257 Hz [42]. Our experiment is based on lead II. Odd numbered records are the training set (called IDS1), and the rest are the testing set (called IDS2).
- MIT-BIH Long-Term ECG Database (MIT-BIH-L) consists of 7 long-term records. The shortest record has 12 hours and the longest record has 22 hours [43]. Our experiment is based on lead ECG1. We separate 90% of the MIT-BIH-L (called LDS1) as the training set, and the remaining data (called LDS2) are the testing set. Table XI shows the heartbeat distributions in the four ECG datasets.
- Wrist **P**hoto**p**lethysmo**g**raphy During Exercise (Wrist-PPG) dataset consists of 19 PPG records of eight participants [56]. Each record contains a series of heartbeat signals corresponding to four exercises taken by the participants. The four exercises are: low resistance exercise bike (low), high resistance exercise bike (high), run on a treadmill (run) and walk on a treadmill (walk). The sampling frequency of each record is 256 Hz. We use overlapping windows to segment the PPG records [57]. A window covers 400 sampling points in a PPG. Each window is shifted by 10 samples. After segmentation, the sample distributions are shown in Table XII. We separate 90% of the PPG segments (called PDS1) as the training set, and the remaining segments (called PDS2) are the testing set.

TABLE XI: Heartbeat Distributions in MIT-BIH-AR, MIT-BIH-SUP, INCART and MIT-BIH-L

Dataset	N	S	V	F	Q
DS1	42328	908	3728	408	8
DS2	42779	1795	4963	387	7
MIT-BIH-SUP	131468	11986	9722	23	78
INCART	119701	1885	16858	208	4
MIT-BIH-L	353644	1497	64058	2904	0

TABLE XII: Action Distributions in Wrist-PPG

Dataset	high	low	run	walk
Wrist-PPG	32752	45218	38101	56903

B. Experiment Setup

Seven experiments are conducted to evaluate the performance of BeatClass. $Experiment_1$ fixes important parameters and compares classification performance of BeatClass with 14 state-of-the-art works [14] [18] [12], [17], [35]–[37] [21]–[23], [32], [38]–[40]. $Experiment_2$ evaluates the performance of MorphGAN. $Experiment_3$ evaluates the classification performance based on different features. $Experiment_4$ evaluates the sustainability of BeatClass based on more physical signal datasets. $Experiment_5$ evaluates the noise-robustness of BeatClass. $Experiment_6$ evaluates the time delay of transferring physical signals over network. We use four indicators to evaluate the classification performance, which are overall accuracy(ACC), precision (PRE), recall (REC) and f1 score (F1). Eqs. (9) to (12) are their calculation methods.

$$ACC = \frac{TP + TN}{TP + TN + FP + FN,} \tag{9}$$

$$PRE = \frac{TP}{TP + FP,} \tag{10}$$

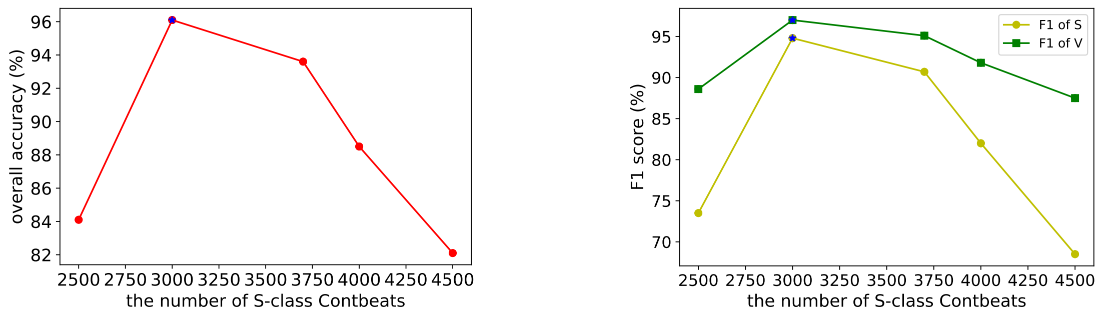
$$REC = \frac{TP}{TP + FN,} \tag{11}$$

$$F1 = \frac{2 * PRE * REC}{PRE + REC.} \tag{12}$$

where, TP, FP, TN and FN are true positives, false positives, true negatives and false negatives, respectively.

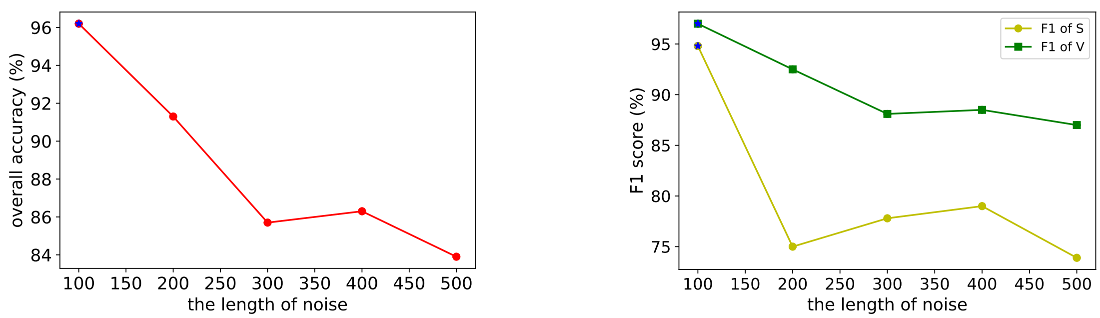
C. $Experiment_1$: Compare Classification Performance of BeatClass with 14 State-of-the-art Works

1) Fixing Important Parameters: This experiment is based on the MIT-BIH-AR dataset. We first test how three important



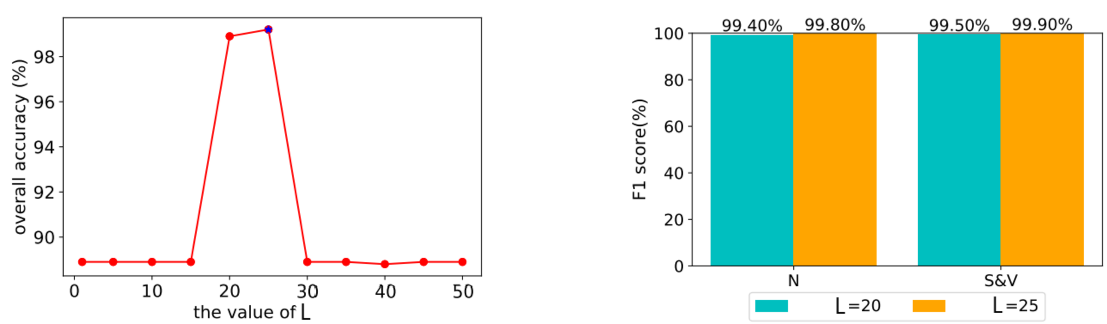
(a) Overall accuracy (b) F1 score

Fig. 6: Performance of BeatClass+MorphGAN with respect to different numbers of S-class Contbeats generated by MorphGAN.



(a) Overall accuracy (b) F1 score

Fig. 7: Performance of BeatClass+MorphGAN with respect to different lengths of noises.



(a) Overall accuracy (b) F1 score

Fig. 8: Performance of BeatClass+MorphGAN with respect to different values of L.

parameters affect the classification performance, which are the number of the augmented S-class heartbeats Num_s, the length of a noise signal $|Znoise|$ and L. The training and testing sets are DS1 and DS2 respectively. As the number of V heartbeats in DS1 is 3728, we estimate that S Contbeats should be augmented to around 3728. Therefore, we set Num_s as 2500, 3000, 3500, 4000 and 4500. The performance of BeatClass with respect to Num_s is shown in Fig. 6. We can see that the best Num_s is 3000.

An important super-parameter of BeatClass+MorphGAN is the length of a noise signal (i.e., $|Z_{noise}|$) that is input into MorphGAN. It can influence the quality of the augmented S Contbeats. We test $|Znoise|$=100, 200, 300, 400 and 500. Fig. 7 shows that the optimal $|Znoise|$ is 100.

The value of L in Rist is also an important super-parameter that affects the performance of BeatClass. We set 11 L values and test the performance of Rist with different L. Fig. 8(a) shows that when L = 20 or 25, BeatClass has the highest ACC. Fig. 8(b) shows that L = 25 is the best value for classifying N and S&V.

In the sequel experiments, we fix Num_s = 3000, $|Znoise|$=100 and L=25.

2) Compare BeatClass with 14 State-of-the-art Works: This experiment compares the performance of BeatClass with

TABLE XIII: Compare BeatClass with State-of-the-art Arrhythmia Detection Work Based on MIT-BIH-AR

Work	Overall ACC (%)	N (%)			S (%)			V (%)		
		REC	PRE	F1	REC	PRE	F1	REC	PRE	F1
BeatClass+MorphGAN	**98.7**	99.9	99.1	**99.5**	94.7	96.8	**94.7**	97.1	99.1	**97.0**
De Chazal et al. (2004) [14]	81.9	86.9	99.2	92.6	75.9	38.5	51.1	77.7	81.9	80.0
Llamedo et al. (2011) [37]	78.0	78.0	99.1	87.3	76.0	41.0	53.3	83.0	88.0	85.4
Ye et al. (2012) [17]	86.4	88.5	97.5	92.8	60.8	52.3	56.3	81.5	63.1	71.2
Zhang et al. (2014) [36]	86.7	88.9	99.0	93.7	79.1	36.0	49.5	85.5	92.8	89.0
Acharya et al. (2017) [18]	71.3	73.3	95.0	82.6	6.3	2.3	3.4	90.8	28.2	43.5
Chen et al. (2017) [35]	93.1	98.4	95.4	96.9	29.5	38.4	33.4	90.8	85.1	77.3
Li et al. (2018) [21]	98.1	99.8	98.0	98.9	68.7	94.7	79.6	95.5	91.1	93.2
Zhai and Tin (2018) [22]	97.6	97.6	98.5	98.0	76.8	74.0	75.4	93.8	92.4	93.1
Niu et al. (2019) [23]	97.5	98.9	97.4	98.1	76.5	76.6	76.5	85.7	94.1	89.7
Mondejar et al. (2019) [38]	96.6	95.9	98.2	97.0	78.1	49.7	60.7	94.7	93.9	94.3
Shi et al. (2019) [39]	95.6	92.0	99.5	95.6	91.7	46.2	61.4	95.1	88.2	91.5
Saadatnejad et al. (2019) [32]	98.4	99.8	97.6	98.7	66.9	95.7	78.7	92.3	98.2	95.2
He et al. (2020) [12]	95.1	97.5	97.6	97.6	83.8	59.4	69.5	80.4	90.2	85.0
Lee et al. (2020) [40]	98.1	97.4	99.6	98.5	91.5	77.6	84.0	89.2	86.0	87.6

TABLE XIV: Performance Evaluation of MorphGAN

Work	Dataset	Euclidean Distance	
		S	V
MorphGAN	MIT-BIH-AR	0.07	0.04
MorphGAN	MIT-BIH-L	0.05	0.14
Deepankar and Rashmi Dutta [58]	MIT-BIH-AR	0.18	0.25

14 state-of-the-art arrhythmia classification models based on DS2 of MIT-BIH-AR. The classification tasks of the 14 works are the same as our task, which are all based on the AAMI standard to classify the five-class heartbeats for different patients. The comparison results are shown in Table XIII. We only compare the classification results of N, S, and V heartbeats, because F and Q heartbeats are like noise and have no classification significance. It is easy to distinguish F and Q heartbeats from N, S and V heartbeats [12].

From Table XIII, the overall ACC of BeatClass is higher than the ACCs of the other work. It is around 17%, 9% and 9% higher than the ACCs of [37], [17] and [36] respectively. For N heartbeats, the classification F1 of BeatClass is around 27%, 13% and 7% higher than the F1's of [18], [37] and [17] respectively. For S heartbeats, the F1 of BeatClass is around 62%, 45%, 44%, and 25% higher than the F1's of [35], [36], [37], and [12] respectively. For V heartbeats, the F1 of BeatClass is around 26%, 17% and 12% higher than [17], [14] and [12] respectively.

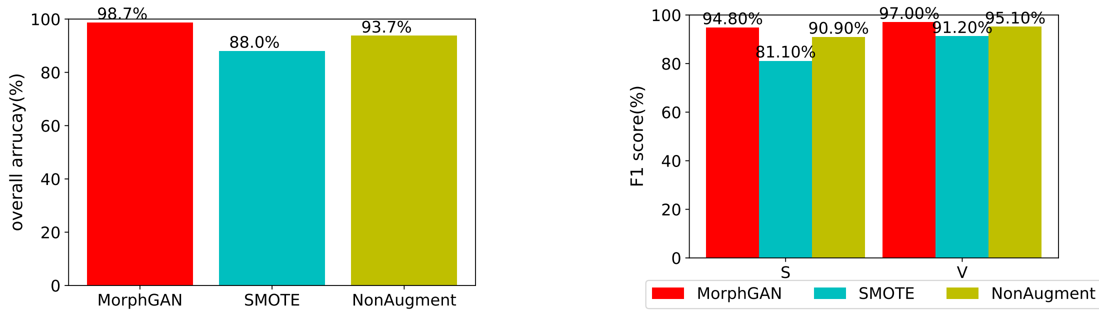
Table XIII shows that many works are difficult to classify S heartbeats. This is because the morphology of the S heartbeat is similar to that of the N heartbeat. Therefore, it is impossible to distinguish S heartbeat from N heartbeat only based on their morphologies. The R-waves of both the S and V heartbeats occur earlier than the R-wave of the N heartbeats. Therefore, Rist in BeatClass first distinguishes N from S and V based on RR sequences. S and V hearbeats are further classified by Morst based on Contbeats and Morphologies. By taking the two-step classification, BeatClass has the best performance compared with the 14 models.

D. $Experiment_2$: Performance of MorphGAN

1) Similarity between the Augmented Data and the Real Data: This experiment evaluates the quality of the data augmented by MorphGAN. The evaluation index is the average Euclidean distance [58] between the augmented data and the real data. Specifically, we augment 1000 S and V heartbeats by using MorphGAN based on two datasets MIT-BIH-AR and MIT-BIH-L. The Euclidean distance is calculated as $\frac{1}{1000\times|DS|}\sum_{i=1}^{|DS|}\sum_{j=1}^{1000}D_{ij}$, where DS represents a dataset, $|DS|$ is the number of heartbeats in DS, and D_{ij} is the Euclidean distance between the ith heartbeat in DS and the jth augmented heartbeat. Table XIV shows the performance comparison between MorphGAN and the data augmentation method in [58]. For both S and V heartbeats, MorphGAN performs better than [58]. For both MIT-BIH-AR and MIT-BIH-L datasets, the augmented S and V heartbeats are similar to the real heartbeats in terms of their Euclidean distances.

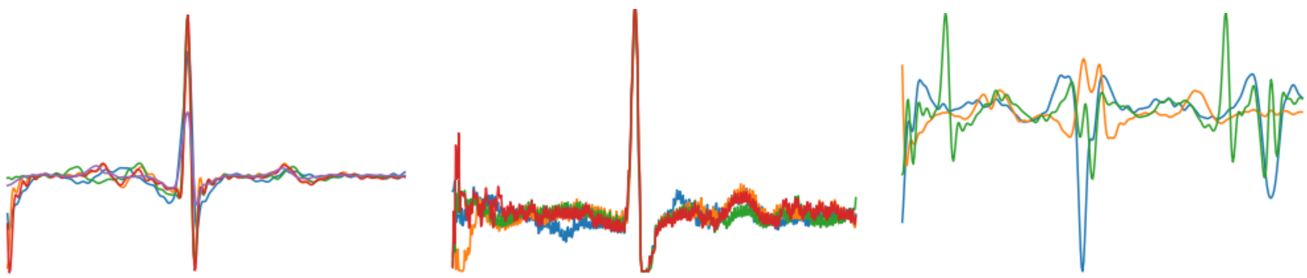
2) Influence of Data Augmentation on Classification Performance: This experiment compares the performance of BeatClass with data augmentation (BeatClass+MorphGAN and BeatClass+SMOTE [24]) and without data augmentation (BeatClass+NonAugment). BeatClass+MorphGAN and BeatClass+SMOTE both augment S Contbeats by using MorphGAN and SMOTE, respectively. S is the most infrequent class compared with N and V in MIT-BIH-AR. From Fig. 9, the ACC of BeatClass+MorphGAN is 4% higher than that of BeatClass+NonAugment. For S and V heartbeats, the F1 scores of BeatClass+MorphGAN are 4% and 2% higher than those of BeatClass+NonAugment respectively.

The ACC of BeatClass+MorphGAN is 10.7% higher than that of BeatClass+SMOTE. For S and V heartbeats, the F1 scores of BeatClass+MorphGAN are 13.7% and 5.8% higher than those of BeatClass+SMOTE respectively. When



(a) Overall accuracy (b) F1 score

Fig. 9: Performance of BeatClass+MorphGAN, BeatClass+SMOTE [24] and BeatClass+NonAugment.



(a) Contbeats from DS2 (b) Contbeats augmented by MorphGAN (c) Contbeats augmented by SMOTE [24]

Fig. 10: Compare the S-class real Contbeats and the Contbeats augmented by MorphGAN and SMOTE.

TABLE XV: Performance of Classification in terms of Different Features Based on MIT-BIH-AR

Work	Overall ACC (%)	N (%)			S (%)			V (%)		
		REC	PRE	F1	REC	PRE	F1	REC	PRE	F1
RR sequence+Contbeat	**98.7**	99.9	99.1	**99.5**	94.7	96.8	**94.7**	97.1	99.1	**97.0**
RR sequence only	92.0	99.9	99.1	99.5	0	-	-	100.0	63.8	77.9
Contbeat only	83.4	90.4	95.5	92.9	8.7	11.1	9.8	85.9	55.4	67.4

TABLE XVI: Performance of BeatClass Based on MIT-BIH-SUP, INCART and MIT-BIH-L

Training Dataset	Overall ACC (%)	N (%)			S (%)			V (%)		
		REC	PRE	F1	REC	PRE	F1	REC	PRE	F1
SDS1	91.4 (SDS2)	99.9	99.4	99.6	87.6	76.1	81.4	75.7	86.2	80.6
IDS1	99.7 (IDS2)	100.0	99.7	99.8	90.2	60.9	72.7	94.2	99.0	96.5
DS1& IDS1	97.8 (DS2)	99.9	99.1	99.5	68.0	95.1	79.3	98.1	84.4	90.7
	99.9 (IDS2)	100.0	100.0	100.0	98.9	99.4	99.4	99.9	99.9	99.9
LDS1	99.2 (LDS2)	100.0	99.2	99.6	77.3	99.2	86.9	99.9	99.5	99.7

we compare the performance of MorphGAN (i.e., BeatClass+MorphGAN) and SMOTE (i.e., BeatClass+SMOTE), the Contbeats of S heartbeats are both augmented based on dataset DS2.

Fig. 10 shows the morphology of some S heartbeats from DS2, the S heartbeats augmented by MorphGAN, and the S heartbeats augmented by SMOTE. We can see that BeatClass+MorphGAN performs much better than BeatClass+SMOTE. SMOTE is more applicable for the featured data oversampling. It is not suitable for augmenting ECG heartbeats that have temporal and sequential relations.

From the two sub-experiments, MorphGAN can augment high quality ECG data.

E. $Experiment_3$: Classification Based on Different Features

This experiment compares the classification performance of BeatClass+MorphGAN based on three different features: RR-sequence+Contbeat, RR-sequence only and Contbeat only. Results are shown in Table XV. Just based on RR sequences, BeatClass misclassifies almost all S heartbeats as V. For S heartbeat classification, the F1 of Beatclass only considering

Contbeat is around 85% lower than the F1 of considering both RR sequences and Contbeats. For V heartbeat classification, the F1 of considering both RR sequences and Contbeats is around 30% higher than the F1 of just considering Contbeats. Overall, the BeatClass considering both RR sequences and Contbeats performs much better than the BeatClass considering RR sequences and Contbeats separately.

F. $Experiment_4$: Sustainability of BeatClass

The experiment validates the sustainability of BeatClass by updating and testing it on five different datasets: MIT-BIH-AR, MIT-BIH-SUP, INCART, MIT-BIH-L, and Wrist-PPG. The ECG recordings in MIT-BIH-SUP, INCART and MIT-BIH-L datasets are first resampled to 360 Hz. We then mix IDS1 and SDS1 and update the BeatClass in $Experiment_3$ based on the mixed dataset. IDS2 and SDS2 are used to test the updated BeatClass.

The testing results are shown in Table XVI. By updating BeatClass based on the mixed datasets DS1 and IDS1, it has good performance for classifying the testing set IDS2. MIT-BIH-SUP does not have lead MLII and lead II ECG signals. This is a major reason why the classification accuracy based on MIT-BIH-SUP is always lower than the classification accuracy based on the other two data sets. The F1 score of the S heartbeat classification on LDS2 does not reach 90%, because the number of S heartbeats in LDS2 is small. The F1 scores of N and V heartbeats based on LDS2 are 99.6% and 96.7%, respectively.

We perform another experiment based on the Wrist-PPG dataset to test the performance of BeatClass for other physical signals with different distributions. BeatClass is fast updated to classify the PPG segments of the four actions. Table XVIII shows the testing results. The overall ACC of BeatClass is around 11% higher than that of the baseline method [57]. The F1 scores of actions high, low, run and walk are all over 85%. They are much higher than the F1 scores of [57].

G. $Experiment_5$: Noise-robustness of BeatClass

There are three main types of noises in ECG signals: baseline wandering noise, myograph interference noise, and power line interference noise [59]. The four ECG datasets used in our experiments all contain the three noises. $Experiments_{1-4}$ show the robustness of BeatClass to the noises. To further evaluate the noise robustness of BeatClass, we add Additive White Gaussian Noise (AWGN) with three different Signal-to-Noise Ratios (SNR) (0db, 5db, 10db) to the MIT-BIH-AR dataset [59]. The classification performance of BeatClass based on such noisy datasets are tested.

The testing results are shown in Table XVIII. Based on the datasets with 0db, 5db and 10db AWGN noises, BeatClass can still achieve good classification performance. The overall ACCs of BeatClass are all over 98%. They are around 23%, 23% and 23% higher than the ACCs of the baseline method [59] based on the datasets with 0db, 5db and 10db AWGNs, respectively. The F1 scores of BeatClass for classifying N, S and V heartbeats also overcome the baseline method [59] dramatically. For example, for S heartbeat, the F1 scores of BeatClass are around 50%, 53% and 57% higher than the F1's of [59] in terms of 0db, 5db and 10db AWGNs, respectively.

H. $Experiment_6$: Time and Space Efficiency Analysis

We test the time delay of sending ECG data to remote Cloud servers in the hospital. The distance from the data sender to the cloud server is about 297.3 kilometers. We send two ECG records of different time lengths (20 and 40 minutes) to the server at 10:00 am, 1:00 am and 8:00 pm from June 10 to June 21, 2021. The sizes of the two ECG records are 1631 and 6521 kilobytes respectively. Table XIX shows the time delay of transferring ECGs over network. ECG data is a single variable time series. It takes little time to transfer a long ECG record over network.

We analyze the time and space complexities of the well-trained Rist and Morst by using the FLOPs and Bytes [60]. FLOPs refer to the total number of operations of a model. It reflects the computing power required by the model. Bytes refers to the number of parameters of a model. It reflects the memory occupied by the model. Table XX shows that the time complexities of Rist and Morst are 0.52 and 18.32 MFLOPs, respectively. Rist and Morst have 109,476 and 6,459,810 parameters and occupy 15.4 and 25.6 MByte (MB) memories, respectively. We can see that the time and space complexities of Rist and Morst are not high.

I. An Engineering Application

BeatClass has significant application value for sustainable development of IoT-based eHealth. Fig. 11 shows an application example of BeatClass [61]. BeatClass is deployed on Cloud servers in hospitals. Patients wear a health bracelet or a portable Holter monitor in daily life. These devices dynamically collect ECG records and transfer them to the BeatClass system via Bluetooth at regular time intervals (e.g., 10∼40 minutes) for classification. If abnormal ECG signals such as V or S are detected, BeatClass will quickly send the ECG records and the detection results to the doctor. The doctor decides the need of treatment based on the type, frequency and duration of abnormality. Patients go to the hospital only if they are asked to. Through remote real-time monitoring and classification of ECG abnormalities, BeatClass can help improve the diagnosis efficiency of doctors and reduce the number of visits to the hospital for patients. It can help doctors discover patient abnormalities on time and reduce patient morbidity. Therefore, BeatClass promotes the sustainable development of IoT-based Telemedicine.

BeatClass is especially useful for dynamic arrhythmia diagnosis based on long-term ECGs [50]. Arrhythmia itself does not cause serious harm to people, but it is often accompanied by some syndromes such as Brugada Syndrome (BrS), Catecholaminergic Polymorphic Ventricular Tachycardia (CPVT), and Sudden Unexplained Death (SUD) (see Fig. 11) [62]. These syndromes can seriously affect people's health. Early diagnosis and timely detection of arrhythmia can effectively prevent the occurrence of syndromes.

Fig. 11 shows an application example of BeatClass [61]. When the heartbeat is classified as S or V, the result will

TABLE XVII: Performance of BeatClass Based on Wrist-PPG

Work	Overall ACC (%)	high (%)			low (%)			run (%)			walk (%)		
		REC	PRE	F1	REC	PRE	F1	REC	PRE	F1	REC	PRE	F1
BeatClass	**88.6**	86.8	91.8	**89.2**	86.3	87.6	**86.9**	87.0	86.2	**86.6**	92.7	89.4	**91.0**
Biagetti et al. (2018) [57]	78.0	87.6	84.9	74.5	55.9	89.9	69.0	97.3	75.9	85.3	76.1	79.4	77.7

TABLE XVIII: Performance of BeatClass Based on MIT-BIH-AR with Different SNR AWGN

Work	Overall ACC (%)	N (%)			S (%)			V (%)		
		REC	PRE	F1	REC	PRE	F1	REC	PRE	F1
AWGN with 0 db										
BeatClass	**98.0**	99.9	99.1	**99.5**	78.6	91.6	**84.6**	95.5	87.4	**91.3**
Singh and Pradhan. (2017) [59]	75.4	76.2	99.0	86.1	50.0	27.1	35.1	79.4	45.2	57.6
AWGN with 5 db										
BeatClass	**98.2**	99.9	99.1	**99.5**	81.5	94.7	**87.7**	97.1	89.3	**93.0**
Singh and Pradhan. (2017) [59]	75.6	76.3	98.9	86.1	49.9	27.2	35.2	79.8	45.4	57.9
AWGN with 10 db										
BeatClass	**98.4**	99.9	99.1	**99.5**	90.1	94.3	**92.1**	96.6	93.8	**95.2**
Singh and Pradhan. (2017) [59]	75.7	76.5	98.9	86.3	49.8	27.1	35.1	80.0	45.6	58.1

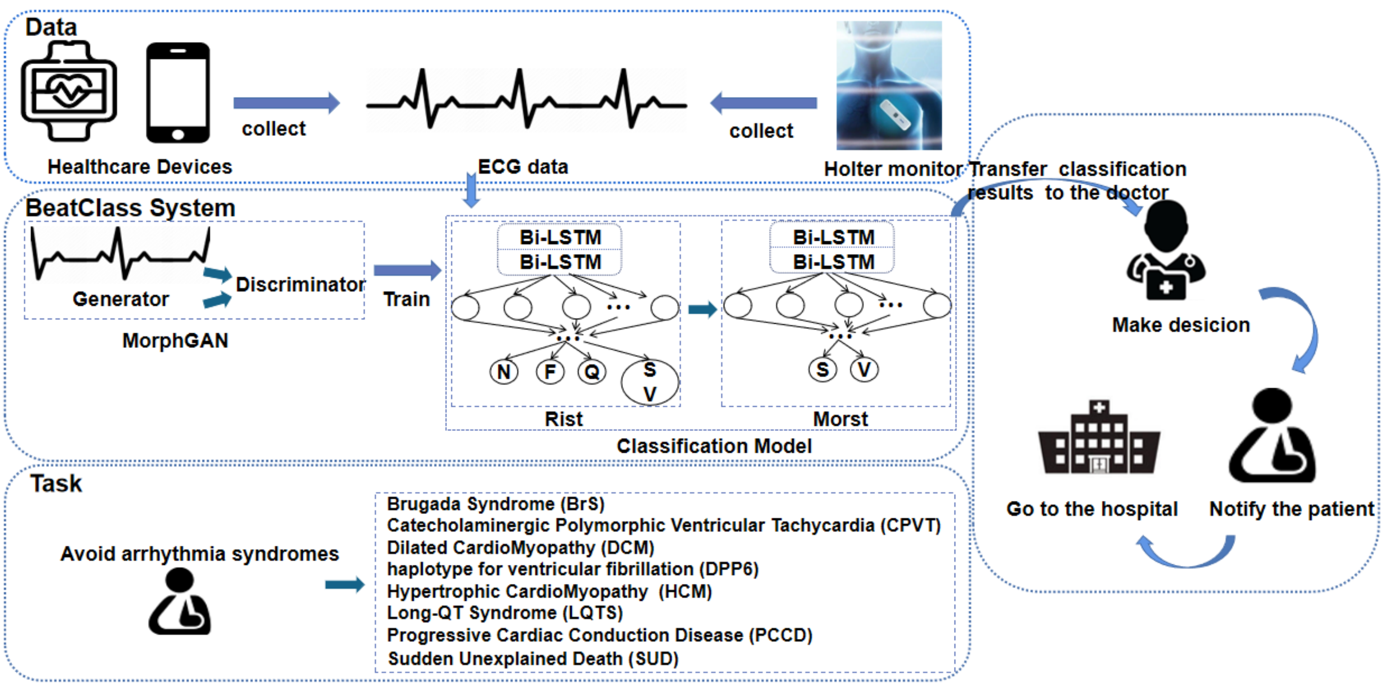


Fig. 11: A sustainable application of BeatClass in IoT-based Telemedicine.

TABLE XIX: Time Delay of Transferring ECG over Network

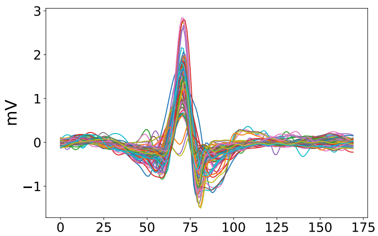
Length of ECG	Average Time Delay (second) of Seven Days		
	8:00 am	1:00 pm	8:00 pm
20 minutes	1	1	1
40 minutes	3	3	3

TABLE XX: The Time and Space Complexities of Rist and Morst

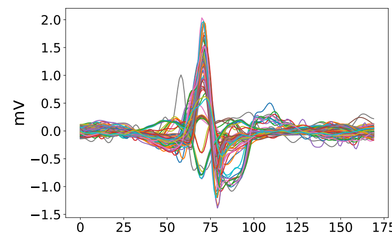
Name	MFLOPs	Bytes (num of parameters)	Model Size (MB)
Rist	0.52	109,476	15.4
Morst	18.32	6,459,810	25.6

be sent to the doctor for further diagnosis. ECGs can be dynamically collected from mobile devices or Holter monitors. BeatClass uses three distinguished features to classify Arrhythmia heartbeats.

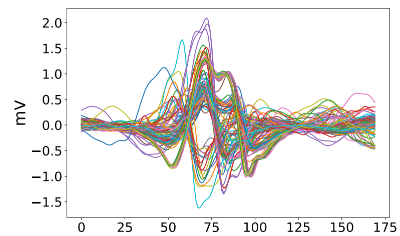
Figs. 12-14 show 100 samples of the three features, from which we can see how they are different and complement with each other. Figs. 12(a)-(e) show the heartbeat morphologies of N, S, V, F, and Q respectively. The morphology of heartbeats



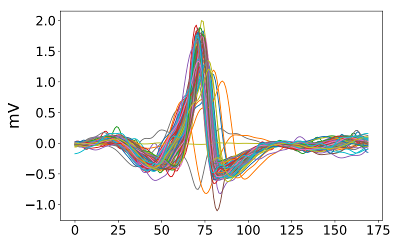
(a) N heartbeats



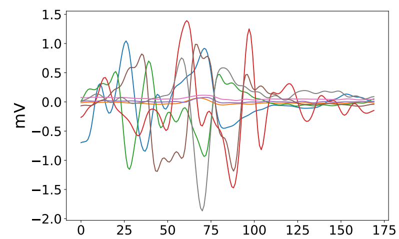
(b) S heartbeats



(c) V heartbeats

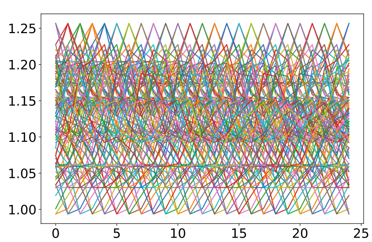


(d) F heartbeats

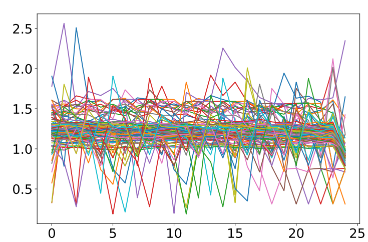


(e) Q heartbeats

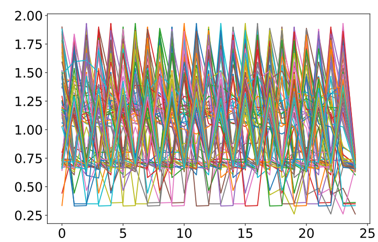
Fig. 12: Examples of morphology of five-class heartbeats.



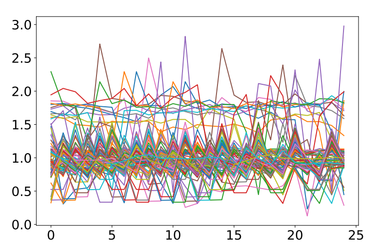
(a) N RRs



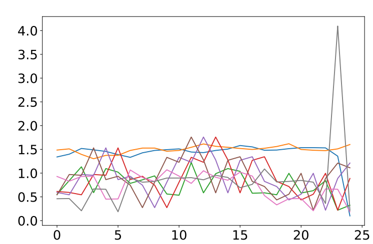
(b) S RRs



(c) V RRs



(d) F RRs



(e) Q RRs

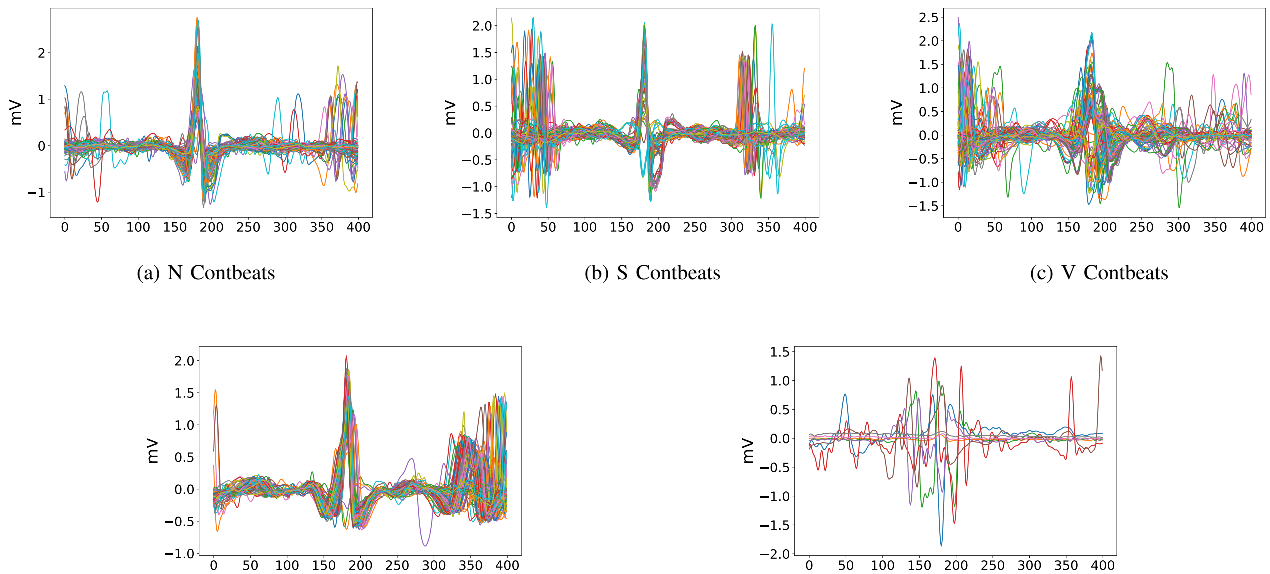
Fig. 13: Examples of five-class RR sequences.

cannot reflect the timing relationship among heartbeats. Figs. 12(a)-(b) show the morphologies of N and S are similar, so it is difficult to distinguish them based on their morphology. Figs. 13(a)-(e) show the RR sequences of N, S, V, F, and Q respectively. We can see that the RR sequences of different types of heartbeats are quite different. Thus, RR sequences are used to detect abnormal heartbeats in BeatClass. Figs. 14 (a)-(e) show the Conbeats of N, S, V, F, and Q respectively. A contbeat can reflect both the morphology and the timing relation between two heartbeats. Contbeats are used to distinguish S and V in BeatClass.

J. Discussion

Experiment results on three real ECG databases demonstrate the effectiveness and robustness of BeatClass:

- **Good performance.** Compared with the 14 state-of-the-art works, BeatClass can classify five-class heartbeats with the best performance. The ACC of BeatClass is



(a) N Contbeats (b) S Contbeats (c) V Contbeats

(d) F Contbeats (e) Q Contbeats

Fig. 14: Examples of five-class Contbeats.

98.7%, which is 27.4% and 0.03% higher than those of the worst and best baseline methods respectively. The F1-scores of classifying N, S&V heartbeats are 99.5%, 94.7% and 97%, which are 0.6%, 10.7% and 1.8% higher than those of the best baseline method respectively.

- **Advanced features.** Based on the feature of RR intervals, N and S&V can be classified with almost 100% accuracy. The F1-scores of BeatClass classifying S and V heartbeats based on RR-sequences and Contbeats achieves 97%, which is 2% higher than the best result in the 14 state-of-the-art works.
- **Effective data augmentation.** The proposed data augmentation model, MorphGAN, can efficiently improve the classification accuracy after augmenting the rare class data. In addition, it performs much better than the classical data augmentation algorithm SMOTE.
- **Sustainability.** BeatClass improves the sustainable development of IoT-based Telemedicine. It can dynamically monitor the health of patients and assist doctors in making diagnostic decisions. It thus reduces the morbidity of patients and saves the medical resources. In addition, the classification capability of BeatClass is sustainable. Based on a small part of a new ECG dataset, a trained BeatClass can be updated to classify the dataset with different sampling rates. This sustainable capability makes BeatClass more robust in the dynamic IoT-based Telemedicine.

V. Conclusion and Future Work

We proposed a two-step heartbeat classification system BeatClass. Rather than using one classification model to classify the five-class heartbeats, BeatClass uses two classification models to improve the classification results step-by-step. The first model Rist classifies heartbeats as N, S, V, F and Q. In this step, N, F and Q can usually be classified accurately. The heartbeats classified as S and V are further classified by Morst. Experiment results demonstrate that using multiple classification models to improve the classification results step-by-step may dramatically enhance the classification performance.

As a physical signal record (e.g., ECG or PPG) is a time series, it is necessary to consider the contextual information and the temporal relation between two heartbeats in heartbeat classification. Therefore, we proposed a new feature Contbeat. It includes the contextual information of a heartbeat and the temporal relationship between the heartbeat and its neighbor heartbeats. By considering such contextual information, the heartbeat classification performance is further improved.

BeatClass may inspire others to use multiple classification models to improve the classification results step-by-step. In addition, considering the contextual and temporal information of heartbeats can help improve the classification performance.

Our experiments are conducted on five public real datasets. We will extend the experiments on more real-time ECG datasets collected from hospitals. Such ECG records may have different baselines and sampling rates. More experiments on these datasets can evaluate the robustness and sustainability of BeatClass more efficiently. We will also integrate the continue learning techniques to improve the sustainability of BeatClass.

Acknowledgment

This work is supported by the National Natural Science Foundation of China (Grants No 61702274) and PAPD.

References

[1] M. Hooshmand, D. Zordan, D. Del Testa, E. Grisan, and M. Rossi, "Boosting the battery life of wearables for health monitoring through the compression of biosignals," *IEEE Internet of Things Journal*, vol. 4, no. 5, pp. 1647–1662, 2017.

[2] U. Satija, B. Ramkumar, and M. Sabarimalai Manikandan, "Real-time signal quality-aware ecg telemetry system for iot-based health care monitoring," *IEEE Internet of Things Journal*, vol. 4, no. 3, pp. 815–823, 2017.

[3] Y. Zeng, N. Xiong, J. H. Park, and G. Zheng, "An emergency-adaptive routing scheme for wireless sensor networks for building fire hazard monitoring," *Sensors*, vol. 10, no. 6, pp. 6128–6148, 2010.

[4] H. Li, J. Liu, R. W. Liu, N. Xiong, K. Wu, and T.-h. Kim, "A dimensionality reduction-based multi-step clustering method for robust vessel trajectory analysis," *Sensors*, vol. 17, no. 8, 2017.

[5] W. Ning, S. Li, D. Wei, L. Z. Guo, and H. Chen, "Automatic detection of congestive heart failure based on a hybrid deep learning algorithm in the internet of medical things," *IEEE Internet of Things Journal*, pp. 1–1, 2020.

[6] B. Lin, F. Zhu, J. Zhang, J. Chen, X. Chen, N. N. Xiong, and J. Lloret Mauri, "A time-driven data placement strategy for a scientific workflow combining edge computing and cloud computing," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 7, pp. 4254–4265, 2019.

[7] W. Fang, X. Yao, X. Zhao, J. Yin, and N. Xiong, "A stochastic control approach to maximize profit on service provisioning for mobile cloudlet platforms," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. PP, no. 99, pp. 1–13, 2016.

[8] Y. Qu and N. Xiong, "Rfh: A resilient, fault-tolerant and high-efficient replication algorithm for distributed cloud storage," in *2012 41st International Conference on Parallel Processing*, 2012, pp. 520–529.

[9] M. Wu, L. Tan, and N. Xiong, "A structure fidelity approach for big data collection in wireless sensor networks," *Sensors*, vol. 15, no. 1, p. 248, 2014.

[10] O. Yildirim, U. B. Baloglu, R.-S. Tan, E. J. Ciaccio, and U. R. Acharya, "A new approach for arrhythmia classification using deep coded features and lstm networks," *Computer Methods and Programs in Biomedicine*, vol. 176, pp. 121–133, 2019.

[11] F. Murat, O. Yildirim, M. Talo, U. B. Baloglu, Y. Demir, and U. R. Acharya, "Application of deep learning techniques for heartbeats detection using ecg signals-analysis and review," *Computers in Biology and Medicine*, vol. 120, p. 103726, 2020.

[12] J. He, J. Rong, L. Sun, H. Wang, and Y. Zhang, "An advanced two-step dnn-based framework for arrhythmia detection," in *Pacific-Asia Conference on Knowledge Discovery and Data Mining*. Springer, 2020, pp. 422–434.

[13] A. for the Advancement of Medical Instrumentation *et al.*, "Testing and reporting performance results of cardiac rhythm and st segment measurement algorithms," *ANSI/AAMI EC57*, 1998-2008.

[14] P. De Chazal, M. O'Dwyer, and R. B. Reilly, "Automatic classification of heartbeats using ecg morphology and heartbeat interval features," *IEEE transactions on biomedical engineering*, vol. 51, no. 7, pp. 1196–1206, 2004.

[15] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, M. Adam, A. Gertych, and R. San Tan, "A deep convolutional neural network model to classify heartbeats," *Computers in Biology and Medicine*, vol. 89, pp. 389–396, 2017.

[16] M. Kachuee, S. Fazeli, and M. Sarrafzadeh, "Ecg heartbeat classification: A deep transferable representation," in *2018 IEEE International Conference on Healthcare Informatics (ICHI)*. IEEE, 2018, pp. 443–444.

[17] C. Ye, B. V. Kumar, and M. T. Coimbra, "Heartbeat classification using morphological and dynamic features of ecg signals," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 10, pp. 2930–2941, 2012.

[18] U. R. Acharya, H. Fujita, O. S. Lih, Y. Hagiwara, J. H. Tan, and M. Adam, "Automated detection of arrhythmias using different intervals of tachycardia ecg segments with convolutional neural network," *Information Sciences*, vol. 405, pp. 81–90, 2017.

[19] K. Jiang, S. Liang, L. Meng, Y. Zhang, P. Wang, and W. Wang, "A two-level attention-based sequence-to-sequence model for accurate inter-patient arrhythmia detection," in *2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, 2020, pp. 1029–1033.

[20] E. J. d. S. Luz, W. R. Schwartz, G. Cámara-Chávez, and D. Menotti, "Ecg-based heartbeat classification for arrhythmia detection: A survey," *Computer Methods and Programs in Biomedicine*, vol. 127, pp. 144–164, 2016.

[21] Y. Li, Y. Pang, J. Wang, and X. Li, "Patient-specific ecg classification by deeper cnn from generic to dedicated," *Neurocomputing*, vol. 314, pp. 336–346, 2018.

[22] X. Zhai and C. Tin, "Automated ecg classification using dual heartbeat coupling based on convolutional neural network," *IEEE Access*, vol. 6, pp. 27 465–27 472, 2018.

[23] J. Niu, Y. Tang, Z. Sun, and W. Zhang, "Inter-patient ecg classification with symbolic representations and multi-perspective convolutional neural networks," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 5, pp. 1321–1332, 2019.

[24] K. N. Rajesh and R. Dhuli, "Classification of imbalanced ecg beats using re-sampling techniques and adaboost ensemble classifier," *Biomedical Signal Processing and Control*, vol. 41, pp. 242–254, 2018.

[25] F. Thabtah, S. Hammoud, F. Kamalov, and A. Gonsalves, "Data imbalance in classification: Experimental evaluation," *Information Sciences*, vol. 513, pp. 429–441, 2020.

[26] H. Zhang, H. Zhang, S. Pirbhulal, W. Wu, and V. H. C. D. Albuquerque, "Active balancing mechanism for imbalanced medical data in deep learning–based classification models," *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, vol. 16, no. 1s, pp. 1–15, 2020.

[27] T. Golany and K. Radinsky, "Pgans: Personalized generative adversarial networks for ecg synthesis to improve patient-specific deep ecg classification," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 01, 2019, pp. 557–564.

[28] C.-C. Lin and C.-M. Yang, "Heartbeat classification using normalized rr intervals and morphological features," *Mathematical Problems in Engineering*, vol. 2014, 2014.

[29] C. C. Lin and C. M. Yang, "Heartbeat classification using normalized rr intervals and wavelet features," in *2014 International Symposium on Computer, Consumer and Control*, 2014, pp. 650–653.

[30] C. Ye, B. V. K. Vijaya Kumar, and M. T. Coimbra, "Heartbeat classification using morphological and dynamic features of ecg signals," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 10, pp. 2930–2941, 2012.

[31] P. de Chazal, M. O'Dwyer, and R. Reilly, "Automatic classification of heartbeats using ecg morphology and heartbeat interval features," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 7, pp. 1196–1206, 2004.

[32] 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, no. 2, pp. 515–523, 2019.

[33] F. Li, Y. Xu, Z. Chen, and Z. Liu, "Automated heartbeat classification using 3-d inputs based on convolutional neural network with multi-fields of view," *IEEE Access*, vol. 7, pp. 76 295–76 304, 2019.

[34] M. Hammad, A. M. Iliyasu, A. Subasi, E. S. L. Ho, and A. A. A. El-Latif, "A multitier deep learning model for arrhythmia detection," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–9, 2021.

[35] S. Chen, W. Hua, Z. Li, J. Li, and X. Gao, "Heartbeat classification using projected and dynamic features of ecg signal," *Biomedical Signal Processing and Control*, vol. 31, pp. 165–173, 2017.

[36] Z. Zhang, J. Dong, X. Luo, K.-S. Choi, and X. Wu, "Heartbeat classification using disease-specific feature selection," *Computers in biology and medicine*, vol. 46, pp. 79–89, 2014.

[37] M. Llamedo and J. P. Martínez, "Heartbeat classification using feature selection driven by database generalization criteria," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 3, pp. 616–625, 2010.

[38] V. Mondéjar-Guerra, J. Novo, J. Rouco, M. G. Penedo, and M. Ortega, "Heartbeat classification fusing temporal and morphological information of ecgs via ensemble of classifiers," *Biomedical Signal Processing and Control*, vol. 47, pp. 41–48, 2019.

[39] H. Shi, H. Wang, Y. Huang, L. Zhao, C. Qin, and C. Liu, "A hierarchical method based on weighted extreme gradient boosting in ecg heartbeat classification," *Computer Methods and Programs in Biomedicine*, vol. 171, pp. 1–10, 2019.

[40] M. Lee, T.-G. Song, and J.-H. Lee, "Heartbeat classification using local transform pattern feature and hybrid neural fuzzy-logic system based on self-organizing map," *Biomedical Signal Processing and Control*, vol. 57, p. 101690, 2020.

[41] S. D. Greenwald, R. S. Patil, and R. G. Mark, "Improved detection and classification of arrhythmias in noise-corrupted electrocardiograms using contextual information," in *[1990] Proceedings Computers in Cardiology*, 1990, pp. 461–464.

[42] A. Goldberger, L. Amaral, L. Glass, J. Hausdorff, P. C. Ivanov, R. Mark, J. Mietus, G. Moody, C. Peng, and H. Stanley, "Components of a new research resource for complex physiologic signals," *PhysioBank, PhysioToolkit, and Physionet*, 2000.

[43] A. L. G. L. H. J. I. P. M. R. M. J. M. G. P. C. Goldberger, A. and H. Stanley, "Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals. circulation [online]," vol. 101, pp. e215–e220, 2000.

[44] H. A. S. M. F. T. Abdelrahman M. Shaker, Manal Tantawi, "Generalization of convolutional neural networks for ecg classification using generative adversarial networks," *IEEE Access*, vol. 8, pp. 35 592–35 605, 2020.
[45] P. Wang, B. Hou, S. Shao, and R. Yan, "Ecg arrhythmias detection using auxiliary classifier generative adversarial network and residual network," *IEEE Access*, vol. 7, pp. 100 910–100 922, 2019.
[46] Z. Zhou, X. Zhai, and C. Tin, "Fully automatic electrocardiogram classification system based on generative adversarial network with auxiliary classifier," *arXiv preprint arXiv:2004.04894*, 2020.
[47] T. Golany, K. Radinsky, and D. Freedman, "Simgans: Simulator-based generative adversarial networks for ecg synthesis to improve deep ecg classification," in *International Conference on Machine Learning*. PMLR, 2020, pp. 3597–3606.
[48] A. Sellami and H. Hwang, "A robust deep convolutional neural network with batch-weighted loss for heartbeat classification," *Expert Systems with Applications*, vol. 122, pp. 75–84, 2019.
[49] B. H. X. C. Bin Zhou, Shenghua Liu and J. Ye, "Beatgan: Anomalous rhythm detection using adversarially generated time series," in *2019 28th International Joint Conference on Artificial Intelligence (IJCAI)*, 2019, pp. 4433–4439.
[50] S. P. Shashikumar, A. J. Shah, G. D. Clifford, and S. Nemati, "Detection of paroxysmal atrial fibrillation using attention-based bidirectional recurrent neural networks," in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. Association for Computing Machinery, 2018, p. 715–723.
[51] X.-C. Cao, B. Yao, and B.-Q. Chen, "Atrial fibrillation detection using an improved multi-scale decomposition enhanced residual convolutional neural network," *IEEE Access*, vol. 7, pp. 89 152–89 161, 2019.
[52] H. Dang, M. Sun, G. Zhang, X. Qi, X. Zhou, and Q. Chang, "A novel deep arrhythmia-diagnosis network for atrial fibrillation classification using electrocardiogram signals," *IEEE Access*, vol. 7, pp. 75 577–75 590, 2019.
[53] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2017.
[54] B. Yi, X. Shen, H. Liu, Z. Zhang, W. Zhang, S. Liu, and N. Xiong, "Deep matrix factorization with implicit feedback embedding for recommendation system," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 8, pp. 4591–4601, 2019.
[55] G. B. Moody and R. G. Mark, "The impact of the mit-bih arrhythmia database," *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 3, pp. 45–50, 2001.
[56] D. Jarchi and A. J. Casson, "Description of a database containing wrist ppg signals recorded during physical exercise with both accelerometer and gyroscope measures of motion," *Data*, vol. 2, no. 1, 2017. [Online]. Available: https://www.mdpi.com/2306-5729/2/1/1
[57] G. Biagetti, P. Crippa, L. Falaschetti, S. Orcioni, and C. Turchetti, "Human activity recognition using accelerometer and photoplethysmographic signals," in *Intelligent Decision Technologies 2017*, I. Czarnowski, R. J. Howlett, and L. C. Jain, Eds. Springer International Publishing, 2018, pp. 53–62.
[58] D. Nankani and R. D. Baruah, "Investigating deep convolution conditional gans for electrocardiogram generation," in *2020 International Joint Conference on Neural Networks (IJCNN)*, 2020, pp. 1–8.
[59] P. Singh and G. Pradhan, "Noise robust automatic heartbeat classification system using support vector machine and conditional spectral moment," *Physical and Engineering Sciences in Medicine*, vol. 43, no. 4, pp. 1387–1398, 2020.
[60] V. Shankar, C. Rowell, W. Hall, A. Mohammadian, M. Schuh, and K. Taylor, "Gigaflop (billion floating point operations per second) performance for computational electromagnetics," *Computing Systems in Engineering*, vol. 3, no. 1, pp. 139–151, 1992.
[61] S. Hong, Y. Zhou, J. Shang, C. Xiao, and J. Sun, "Opportunities and challenges of deep learning methods for electrocardiogram data: A systematic review," *Computers in Biology and Medicine*, p. 103801, 2020.
[62] N. Hofman, H. L. Tan, M. Alders, I. Kolder, S. de Haij, M. M. Mannens, M. P. Lombardi, R. H. Lekanne dit Deprez, I. van Langen, and A. A. Wilde, "Yield of molecular and clinical testing for arrhythmia syndromes: report of 15 years' experience," *Circulation*, vol. 128, no. 14, pp. 1513–1521, 2013.

Le Sun Dr. Le Sun is a Professor in the School of Computer and Software, Nanjing University of Information Science and Technology, China. She received her PhD degree in 2016, from Victoria University, Australia, and received the Chancellor-citation Best PhD Award. In 2018, she was a visiting scholar of Michigan State University. She has published over 30 papers on high quality conferences and journals, e.g. IEEE Transactions on Services Computing, ACM Transactions on Internet Technology, IEEE Access, and Future Generation Computer Systems. Her research interests include Medical data stream mining, services computing, and fog computing.

Yilin Wang Miss Yilin Wang is a master student in the School of Computer and Software, Nanjing University of Information Science and Technology, China. She received her bachelor degree from Nanjing University of Information Science and Technology, China, in 2019. Her research interests include smart medical and data mining.

Zhiguo Qu Dr. Zhiguo Qu received the Ph.D. degree in information security from Beijing University of Posts and Telecommunications, China, in 2011. From 2012 to 2014, he worked as a Post-doc Researcher Fellow at Dublin City University in Ireland. In July 2011, he joined Nanjing University of Information and Technology in China, where he is currently an Associate Professor in the School of Computer and Software. His research interests include quantum secure communication, quantum information hiding, quantum computing, data mining and digital watermarking.

Neal N. Xiong Dr. Neal N. Xiong (S'05–M'08–SM'12) is current a Professor at Department of Computer Science and Mathematics, Sul Ross State University, Alpine, USA. He received his both PhD degrees in Wuhan University (2007, about sensor system engineering), and Japan Advanced Institute of Science and Technology (2008, about dependable communication networks), respectively. Before he attended Sul Ross State University, he worked in Northeastern State University, Georgia State University, Wentworth Technology Institution, and Colorado Technical University (full professor about 5 years) about 16 years. His research interests include Cloud Computing, Security and Dependability, Parallel and Distributed Computing, Networks, and Optimization Theory. He published over 200 international journal papers and over 100 international conference papers. Some of his works were published in IEEE JSAC, IEEE or ACM transactions, ACM Sigcomm workshop, IEEE INFOCOM, ICDCS, and IPDPS. He has been a General Chair, Program Chair, Publicity Chair, Program Committee member and Organizing Committee member of over 100 international conferences, and as a reviewer of about 100 international journals, including IEEE JSAC, IEEE SMC (Park: A/B/C), IEEE Transactions on Communications, IEEE Transactions on Mobile Computing, IEEE Trans. on Parallel and Distributed Systems. He is serving as an Editor-in-Chief, Associate editor or Editor member for over 10 international journals (including Associate Editor for IEEE Tran. on Systems, Man & Cybernetics: Systems, Associate Editor for IEEE Tran. on Network Science and Engineering, Associate Editor for Information Science, Editor-in-Chief for Journal of Internet Technology (JIT), and Editor-in-Chief for Journal of Parallel & Cloud Computing (PCC)), and a guest editor for over 10 international journals, including Sensor Journal, WINET and MONET. He has received the Best Paper Award in the 10th IEEE International Conference on High Performance Computing and Communications (HPCC-08) and the Best student Paper Award in the 28th North American Fuzzy Information Processing Society Annual Conference (NAFIPS2009).