IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 30 July 2023, accepted 7 August 2023, date of publication 15 August 2023, date of current version 18 August 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3305473

RESEARCH ARTICLE

Heartbeat Dynamics: A Novel Efficient Interpretable Feature for Arrhythmias Classification

XUNDE DONG[1] **AND WENJIE SI**[2]

[1]School of Automation Science and Engineering, South China University of Technology, Guangzhou 510641, China
[2]School of Electrical and Control Engineering, Henan University of Urban Construction, Pingdingshan 467036, China

Corresponding author: Xunde Dong (audxd@scut.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 62003141 and Grant U20A20224; in part by the Natural Science Foundation of Guangdong Province, China, under Grant 2021A1515011598; and in part by the Fundamental Research Funds for the Central Universities, under Grant 2022ZYGXZR023.

ABSTRACT Arrhythmias are a significant class of cardiovascular diseases, and timely and accurate detection is critical in preventing high-risk events such as sudden cardiac death. Despite the attention that automatic detection of arrhythmias based on electrocardiogram (ECG) has received, static features used in traditional methods fail to adequately describe the various weak changes of the ECG, resulting in significant but weak pathological information being overlooked. Although deep learning (DL) extracted features demonstrate efficiency in arrhythmia classification, the interpretability of DL methods remains challenging. In this study, we propose a novel and efficient interpretable feature for arrhythmia classification, heartbeat dynamics. It models morphological changes in the heartbeat, and is more sensitive to weak heartbeat variations, and reflects underlying dynamical changes throughout the cardiac cycle at the electrophysiological level. To evaluate its efficiency for arrhythmia classification, we conducted experiments on the MIT-BIH arrhythmia database, using three classical classifiers: k-nearest neighbor (KNN), random forest (RF), and support vector machine (SVM). Our proposed method achieves 99.41% accuracy, 99.10% precision, 98.84% recall, and 0.9897 F1 score with KNN as the classifier, comparable to or better than most DL-based methods. These results indicate that heartbeat dynamics has a strong ability to discriminate between different classes of heartbeats. We anticipate that the heartbeat dynamics feature will enhance the generalization capacity of the arrhythmia detection algorithm when integrated with other static features.

INDEX TERMS ECG, heartbeat dynamics, arrhythmia detection, deterministic learning.

I. INTRODUCTION

Cardiovascular disease is a leading disease that threatens human life and health worldwide. The annual report on cardiovascular health and diseases in China (2021) shows that the prevalence of cardiovascular disease in China is rising, with a projected 330 million people suffering from cardiovascular disease [1]. More importantly, cardiovascular disease is the leading cause of death in the population, accounting for 2 out of every 5 deaths. Statistics show that more than 80% of cardiovascular patients suffer from arrhythmias, which can easily lead to stroke, sudden cardiac death and other malignant events. Therefore, timely and rapid detection of arrhythmias and intervention before serious events occur is critical to prevent high-risk events such as heart attack, stroke and sudden cardiac death.

As an electrophysiological signal capable of characterizing the state of the heart, ECG signals are essential for the detection and diagnosis of cardiovascular diseases. Because ECG is non-invasive, cost-effective and simple, it is the most commonly used clinical tool to target arrhythmias. However, this process requires cardiologists to observe long-term ECG recordings, which are undoubtedly subjective, time-consuming and labor-intensive. Therefore, computer-aided

The associate editor coordinating the review of this manuscript and approving it for publication was Ines Domingues.

automatic ECG signal classification systems have emerged as an essential aid for arrhythmia analysis.

Machine learning, the most active research area of artificial intelligence, has been widely and successfully applied to the analysis of medical signals such as ECG and electroencephalogram [2], [3]. In the field of ECG analysis, researchers have proposed various methods for automatic detection of cardiac arrhythmias using traditional machine learning approaches. Typically, these methods involve three main steps: signal preprocessing, feature extraction, and pattern recognition and classification. Feature extraction plays a crucial role in improving the performance of heartbeat classification. Popular feature extraction techniques include principal component analysis (PCA) [4], [5], [6], wavelet transform (WT) [7], [8], [9], discrete wavelet transform (DWT) [10], [11], [12], independent component analysis (ICA) [13], [14], and additional hand-crafted features [15], [16], [17], [18], [19], [20], [21], [22]. For example, Martis et al. [5], [6] used PCA for feature extraction and compression in ECG analysis. Abdelazez et al. [9] used WT for feature extraction from compressed ECG signals and constructed an AF detector using random forests. Rai et al. [10] introduced another feature extraction approach combining DWT with morphology and achieved good performance when combined with an artificial neural network (ANN) classifier. Specially, WT has been considered particularly suitable for ECG signal analysis due to the inherent non-stationarity of ECG signals [23]. Li et al. [13] proposed a novel multi-domain feature extraction method combining kernel ICA and WT, followed by PCA for feature compression. In [14], ICA was employed for the purpose of dimensionality reduction and feature extraction of the ECG signal Hand-crafted features are usually selected through experience. Those features usually include entropy-based features [15], linear and nonlinear features [16], higher order spectra [17], [18], [19], higher order statistics [20], [24], statistical features [21], [22], and sparse features [25].

Once features have been extracted from ECG arrhythmia signals, classification models can be developed to identify different types of arrhythmia. Support vector machines (SVM) [26], [27], [28], [29], [30] and ANN [10], [31] are two popular algorithms commonly used for classification. For example, Mondejar et al. [29] proposed a method for automatic classification of ECG signals by combining several SVMs. Geweid et al. [30] used a hybrid approach with dual SVMs to detect atrial fibrillation. Multilayer perceptron (MLP) is the most widely used ANN architecture for arrhythmia classification. Rai et al. [10] used back-propagation neural networks, feed-forward networks and MLPs to classify ECG signals into normal and abnormal classes. In another study [31], arrhythmias were detected by using feature vectors as input to the MLP model.

The remarkable success of DL in machine vision, image recognition, and speech recognition has prompted extensive research exploring its potential in arrhythmia classification [32], [33], [34], [35], [36], [37], [38], [39], [40], [41]. For instance, Hannun et al. [32] developed a deep neural network (DNN) capable of classifying 12 different rhythm classes. Khan et al. [34] implemented a long short-term memory network (LSTM) DL approach for arrhythmia classification. In [33], a 1D-convolutional neural network (CNN) model was proposed for the automatic classification of cardiac arrhythmia. Li et al. [37] proposed a 2D-CNN model that employs overall feature maps of heartbeats generated through empirical modal decomposition for arrhythmia classification. Recurrent neural networks (RNNs) were employed for differentiating between normal and abnormal heartbeats in [39], whereas [41] presented an improved deep residual network for arrhythmia classification. Additionally, DenseNet [42] and ResNet [43] have also been employed in the classification of arrhythmias.

DL-based approaches integrate feature extraction and classification as a whole, rather than explicitly describing them as two separate modules. DL models can automatically identify the features needed for classification by using large amounts of ECG data. However, the automatic learning of features from raw data rather than doing it manually makes the interpretability of DL methods still a general challenge, which is especially critical for medical applications, as the mysterious process may not be accepted by medical professionals [44]. More importantly, although researchers have continued to use various advanced DL models for arrhythmia classification, there has been no qualitative improvement in classification performance. This suggests that DL-based methods may have reached their limits. Moreover, DL-based approaches suffer from complex network structures with a large number of parameters, which limits their deployment to wearable devices with limited memory and power consumption.

Since it is difficult to make further improvements in arrhythmia classification by static feature-based machine learning methods and DL-based methods, it is necessary to study it from a different perspective. Indeed, ECG is a class of morphologically variable temporal patterns that are not adequately described by static features alone. The ECG dynamics proposed in the literature [45] is a fresh perspective characterization of the ECG that provides a more comprehensive description of the ECG than static features. It is the accurate modeling of the various morphological changes in the ECG that reflect the depolarization and repolarization processes of cardiomyocytes [46], [47], [48], which makes it interpretable.

The main purpose of this paper is to evaluate the ability of heartbeat dynamics to discriminate between different classes of heartbeats. It is expected that the heartbeat dynamics feature will enhance the generalization capacity of the arrhythmia detection algorithm when integrated with other static features. We first preprocess the single-lead ECG with denoising, R-peak detection, and heartbeat segmentation. Then deterministic learning (DetL) is applied to extract heartbeat dynamics. To reduce the consumption of computational resources, we use principal component analysis (PCA) methods to reduce the dimensionality of the extracted heartbeat dynamics and obtain low-dimensional feature vectors for classification. In the classification stage, we use three simple and classical classifiers, KNN, SVM and RF to classify heartbeats.

In contrast to other machine learning based approaches, the highlights of this study can be summarized as follows.

- Heartbeat dynamics is a novel way to describe ECG more comprehensively and is used as an exclusive classification feature.
- It is interpretable compared to DL-extracted features as it is the accurate modeling of various morphological changes in the ECG.
- It has good discriminating for different classes of heartbeats, making it easy to deploy into wearable devices with simple classifiers.

The rest of the paper is organized as follows. Section II is devoted to the materials and methods, including the databases and techniques used in this paper. Section III presents the experimental results. In Section IV, we discuss and analyze the experimental results. A conclusion is given in Section V.

II. MATERIALS AND METHODS

A. MATERIALS

1) MIT-BIH ARRHYTHMIA DATABASE

This study evaluated the classification performance of the proposed method on the benchmark MIT-BIH arrhythmia database [49]. It was created in 1980 to motivate the development of techniques for automatic detection and classification of arrhythmia. 48 ECG recordings taken from 47 individuals are available in this database. Each recording contains two-lead (40 recordings: modified lead II and lead V1, 2 recordings: modified lead II and lead V5; 2 recordings: modified lead II and lead V2; 2 recordings: lead V5 and lead V2; 1 recording: lead V5 and modified lead II; 1 recording: modified lead II and V4) 30-minute ECG signals. In the experimental section, heartbeat classification experiments based on single-lead ECGs will be performed on each lead ECG separately, and the two leads will be denoted as lead 1 and lead 2 in the following for simplicity. The ECG recordings were independently annotated by at least two cardiologists with temporal information and heartbeat classes, and the entire database has approximately 109000 manually annotated heartbeat labels.

2) DETERMINISTIC LEARNING

The deterministic learning (DetL) theory was introduced in 2007 by Wang and Hill [50] and focuses on problems such as accurate identification and rapid temporal pattern recognition. In DetL theory, temporal patterns are time-varying regression trajectories typically generated by the following dynamical system:

$$\dot{y} = F(y;p),\ y(t_0) = y_0 \tag{1}$$

where $y = [y_1, \ldots, y_n]^T \in R^n$ is the system state with regression characteristics, p is a unknown parameters vector, and $F(y;p) = [f_1(y;p), \ldots, f_n(y;p)]^T$ is the unknown system dynamics.

From the dynamical system equation (1), it can be seen that the system state y is completely determined by the system dynamics $F(y;p)$. In other words, the dynamics of the system is the most essential characteristic of such a dynamical system. To achieve an accurate identification of the unknown system dynamics $F(y;p)$, we use the following estimation system:

$$\dot{\hat{y}}_i = -c_i(\hat{y}_i - y_i) + \hat{W}_i S_i(y), \tag{2}$$

where $\hat{y}_i$ is the estimator state, y_i is the system state of the dynamical system (1), $c_i > 0$ is the design constant, $\hat{W}_i S_i(y)$ is the RBF network for approximating the unknown dynamics $f_i(y;p)$ of the dynamical system (1), with $\hat{W}_i = [\hat{w}_{i1}, \ldots, \hat{w}_{iN}]^T \in R^N$ and $S_i(y) = [s_{i1}(\| y - \xi_1 \|), \ldots, s_{iN}(\| y - \xi_N \|)]^T$, $s_{ij}(\cdot)$ is the Gaussian function, ξ_j $(j = 1, \ldots, N)$ are distinct centers.

The weight estimate $\hat{W}_i$ is updated according to the following law:

$$\dot{\hat{W}}_i = -\Gamma_i S_i(y)\tilde{y}_i - \sigma_i \Gamma_i \hat{W}_i, \tag{3}$$

where $\Gamma_i = \Gamma_i^T > 0$, $\tilde{y}_i = \hat{y}_i - y_i$ and $\sigma_i > 0$ is a small constant.

DetL theory has shown unknown dynamics $f_i(y;p)$ can be accurately identified along almost all temporal patterns with regression trajectories [50], [51], [52] and expressed as the following equation:

$$f_i(y;p) = \bar{W}_i S_i(y) + \epsilon_i, \tag{4}$$

where $\bar{W}_i$ is a constant weight vector, and ϵ_i is the approximation error. Thus, the system dynamics $F(y;p)$ can be accurately modeled and expressed as follows:

$$F(y;p) = \bar{W}S(y) + \epsilon \tag{5}$$

where $\bar{W}S(y) = [\bar{W}_1 S_1(y), \cdots, \bar{W}_n S_n(y)]^T$, $\epsilon = [\epsilon_1, \cdots, \epsilon_n]^T$. That is, the system dynamics is accurately identified and stored in $\bar{W}S(y)$, which can be directly used for classification and recognition of temporal patterns.

B. METHODS

In this subsection, we outline the proposed method in three main steps. The first step is to preprocess the ECG signal, which involves both denoising and segmenting the heartbeat. Second, DetL theory is used to extract heartbeat dynamics. Third, the extracted heartbeat dynamics are used as input to KNN, SVM, and RF classifiers for heartbeat classification.

1) PREPROCESSING

The ECG is susceptible to contamination by various noises during the measurement. Removing such noise is crucial for accurate recognition and classification. We preprocessed the ECG by using the approach proposed by [53].

All ECG signals in the MIT-BIH database were first preprocessed using a 200-ms width median filter to remove P-wave and QRS complexes, followed by a 600-ms width median filter to remove T-wave. The resulting signal is then treated as a baseline, which is subsequently subtracted from the raw ECG to yield the baseline corrected ECG. A 12th-order FIR low-pass filter with a cutoff frequency of 35 Hz was then used to remove power lines and high-frequency noise from the baseline corrected ECG.

In this paper, a window of -300 ms to 400 ms around the R-peak is chosen as a heartbeat, where the location of the R-peak is provided by the annotation files of the MIT-BIH arrhythmia database. The sampling frequency of the MIT-BIH database is 360 HZ, so a heartbeat contains 252 sampling points. Additionally, each heartbeat is normalized to zero mean and unit variance by subtracting the mean and dividing by the standard deviation.

2) HEARTBEAT DYNAMICS EXTRACTION

The ECG signal is a comprehensive representation of the cardiac electrical conduction system at the surface of the body and is a quasi-periodic time-varying signal, which enables its accurate identification using DetL theory. To describe the various changes in the ECG, the most commonly used features are the morphological features of the ECG. Morphological features essentially reflect the depolarization and repolarization processes of cardiomyocytes, such as changes in the QRS waves that reflect changes in the cardiomyocyte depolarization and conduction of signals to the rest of the heart and changes in the ST segments that reveal abnormalities in the cardiomyocyte repolarization process. Heartbeat dynamics is the accurate modeling of various morphological changes in the ECG. In this way, it is interpretable. It essentially models the pattern of electrical activity associated with myocardial electrical excitation, characterizing the rate and amplitude of myocardial electrical excitation. In contrast to traditional morphological features, heartbeat dynamics provides a more comprehensive and detailed description of cardiac electrical activity and is more sensitive to various small morphological changes in the heartbeat.

To extract the dynamics of the heartbeat, we formulate the complex, high-dimensional, continuous nonlinear dynamical system of the cardiac electrical conduction system in the following form:

$$\dot{E}(t) = F(E(t); p) \tag{6}$$

where $E(t) = [e_1(t), \cdots, e_n(t)]^T$ is the ECG signal measured at the body surface, $F(E(t)) = [f_1(E(t)), \cdots, f_n(E(t))]^T$ is the system dynamics, which uniquely determines the morphology of the ECG signal, and n is the number of leads, p is the parameter vector. For a heartbeat of a single-lead ECG signal (i.e., $n = 1$), the dynamical system that generated it can be expressed as follows:

$$\dot{e}_H(t) = f(e_H(t); p_e) \tag{7}$$

where $e_H(t)$ is a heartbeat of the single-lead ECG signal, $f(e_H(t); p_e)$ is the unknown system dynamics, and p_e is a unknown parameter vector.

To extract the dynamics of heartbeat $e_H(t)$, we first construct the following estimation system:

$$\dot{\hat{e}}_H(t) = -c(\hat{e}_H(t) - e_H(t)) + \hat{W}S(e_H(t)) \tag{8}$$

where $\hat{e}_H(t)$ is the estimation of heartbeat $e_H(t)$, $\hat{W}S(e_H(t))$ is the RBF network for approximating the unknown dynamics $f(e_H(t); p_e)$ and $c = 5$.

The weight estimate $\hat{W}$ is updated using the following law:

$$\dot{\hat{W}} = -\Gamma S(e_H(t))\tilde{e}_H(t) - \sigma\Gamma\hat{W}, \tag{9}$$

where $\Gamma = 20$, $\tilde{e}_H(t) = \hat{e}_H(t) - e_H(t)$ and $\sigma = 0.01$.

According to the DetL theory, the heartbeat dynamics $f(e_H(t))$ of a single-lead ECG can be accurately modeled and expressed in the following form:

$$f(e_H(t)) = \bar{W}S(e_H(t)) + \epsilon \tag{10}$$

where $\bar{W}$ is the constant weight of RBF network and ϵ is the modeling error that can be arbitrarily small.

According to equation (7), it can be observed that the waveform of $e_H(t)$ is completely determined by the heartbeat dynamics, denoted as $f(e_H(t))$. In contrast, the various features extracted by existing methods, including linear features, nonlinear features and features extracted by DNN, are all derived from the heartbeat signal $e_H(t)$. As a result, heartbeat dynamics provide a more fundamental description than the various features extracted by existing methods.

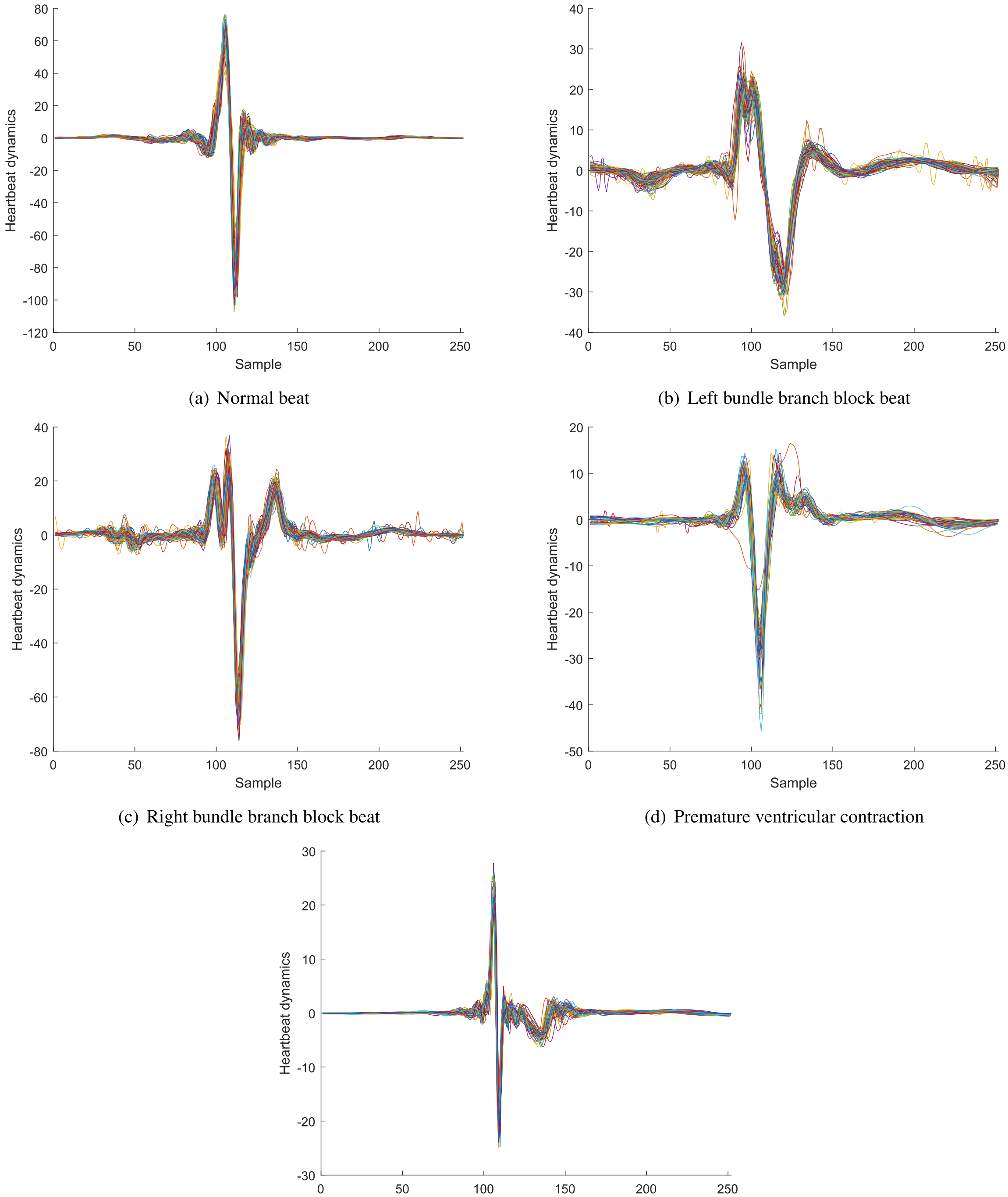
The extracted heartbeat dynamics has the same dimension as the heartbeat signal, which is also a 1×252 feature vector. Figure 1 shows the heartbeat dynamics of the five classes extracted using DetL theory. It visually demonstrates that the dynamics of different classes of heartbeats are highly discriminative.

3) PRINCIPAL COMPONENT ANALYSIS

As mentioned above, the extracted heartbeat dynamics is a 1×252 feature vector. PCA is used to reduce the dimensionality of heartbeat dynamics to reduce the computational resource consumption. It is a mathematical algorithm that reduces the dimensionality of the data while retaining most of the variation in the dataset. It accomplishes this reduction by identifying directions, called principal components, along which the variation in the data is maximal [54]. The computation of principal components involves calculating the covariance matrix of the data, decomposing it into eigenvalues, sorting the eigenvectors in decreasing order of eigenvalues, and finally projecting the data into a new basis defined by principal components by performing an inner product of the original signal and the sorted eigenvectors.

4) CLASSIFIERS

(1) **K-Nearest Neighbor**: The KNN algorithm is one of the simplest methods of all machine learning algorithms



(a) Normal beat

(b) Left bundle branch block beat

(c) Right bundle branch block beat

(d) Premature ventricular contraction

(e) Paced beat

FIGURE 1. **The heartbeat dynamics of five classes heartbeat.**

that can be used for classification and regression problems. It classifies a sample by estimating the majority vote of its neighbors. Since it is a non-parametric algorithm, which makes it very easy to implement, it remains one of the most popular algorithms for pattern classification.

A formal notation for KNN in classification problems is as follows. Let X_{tr} and X_{te} be the training and test sets, respectively. Each sample x_i is a vector $(x_i^1, x_i^2, \ldots, x_i^D)$, where $x_i^j (j = 1, 2, \cdots, D)$ is the value of the j-th feature of the i-th sample. The class of each sample in X_{tr} is known, while the class of each sample in X_{te} is unknown. For a sample x_t in X_{te}, the KNN algorithm calculates the distance between this sample and all samples of X_{tr}, and selects the k closest samples in X_{tr} to be sorted by distance from the largest to the smallest. The class of x_t is then determined by voting on the class labels of its k nearest neighbors.

(2) **Support vector machine**: SVM is a generalized linear classifier that performs binary classification in a supervised manner. The main objective of SVM is to determine an optimal hyperplane to separate two classes of samples. The optimal hyperplane refers to the decision boundary that achieves the minimum misclassification error during the training phase. The determination of the decision boundary is transformed into a convex quadratic optimization solution problem. To maximize the separation, the SVM uses the part of the training samples closest to the optimal decision boundary as the support vector. For nonlinear classification problems, Cortes and Vapnik [55] introduced soft marginal and kernel trick methods to address the limitations of linear SVMs. In soft-margin SVM, slack variables are added to handle nonlinearly separable data. In kernel SVM, kernel functions are used to map samples from a low-dimensional input space to a high-dimensional space such that samples are linearly separable in the new feature space.

(3) **Random forest**: RF is an ensemble learning method developed by Breiman [56] for solving classification and regression problems. Ensemble learning is a machine learning scheme that improves accuracy by integrating multiple models to solve the same problem. In other words, the integration of multiple classifiers reduces the classification error, especially in the case of single-classifier instabilities. Based on the classification results of multiple classifiers, a voting scheme is designed to assign a label to the samples to be classified. The commonly used voting method is majority voting due to its simplicity and effectiveness. It assigns a label to each unlabeled sample with the highest number of votes.

Boosting and bagging are two widely used types of ensemble learning methods. Boosting is the process of constructing a sequence of models, each of which attempts to correct the errors of the previous model in that sequence. Bagging is the best-known representative of parallel integration learning methods [57] and aims to improve the stability and accuracy of integrated models while reducing variance. RF was the first successful Breiman's bagging sampling method that combined bagging, random decision forests, and feature random selection. It has been widely adopted and applied as a standard classifier to a variety of prediction and classification tasks, such as those in bioinformatics [58], computer vision [59].

III. EXPERIMENTAL

In this section, we conduct experiments based on the MIT-BIH arrythmia database. All ECG recordings from the MIT-BIH arrythmia database were considered. It is important to note that although the database contains 15 classes of heartbeats, there is a severe imbalance in the distribution of instances across these classes. In particular, five classes of heartbeats have more than 7000 samples each, with the normal beat class having more than 70000 samples. In contrast, some classes of heartbeats contain only a few dozen or even a few samples, posing a challenge in capturing discriminative features for accurate classification of these minority heartbeat classes. Therefore, following the approach of most existing arrhythmia classification studies, this study focuses on the five classes with the highest number of samples for classification experiments. The classes and their numbers are as follows: normal beat: 75020; left bundle branch block (LBBB) beat: 8072; right bundle branch block (RBBB) beat: 7255; premature ventricular contraction (PVC): 7124; paced beat (PB): 7130. To more adequately demonstrate the discriminative power of heartbeat dynamics, we conduct experiments on the original dataset, the balanced datasets after undersampling and oversampling, respectively.

The following four performance metrics: precision, recall, accuracy and F1-score are selected to evaluate the proposed method, and the corresponding formulas are given as follows:

$$\mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}}$$
$$\mathrm{Recall} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}$$
$$\mathrm{Accuracy} = \frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN}}$$
$$\mathrm{F1-score} = \frac{2\times\mathrm{Precision}\times\mathrm{Recall}}{\mathrm{Recall}+\mathrm{Precision}}$$

where TP represents the true positive counts, TN represents the true negative counts, FP represents the false positive counts and FN represents the false negative counts.

A. EXPERIMENTS ON ORIGINAL HEARTBEAT DATASET

In this subsection, we conduct experiments on the original heartbeat dataset, which is an imbalanced dataset, to estimate the classification performance using heartbeat dynamics as the exclusive classification feature. Since the heartbeat dynamics is a 1×252 vector, we use PCA to reduce the dimensionality of the heartbeat dynamics to improve the real-time performance of the algorithm. As the first 15 principal components already contain about 92% information of the heartbeat dynamics, the heartbeat dynamics is reduced as a 1×15 feature vector, which can not affect the classification performance and can also avoid the overfitting problem during training. In the text below, the classification feature used in all experiments is the 1×15 feature vector after dimensionality reduction of the heartbeat dynamics and are not described anymore.

In the following experiments, we choose three classical and simple classifiers, KNN, SVM and RF, to validate the effectiveness of heartbeat dynamics for heartbeat classification. The three classifiers are direct calls from the machine learning library scikit-learn, and all arguments are default without tuning and optimization. To give the reader a more intuitive understanding of these parameters, we present the default values of some main parameters for the three classifiers. KNN: n_neighbours=5, weights='uniform', algorithm='auto', leaf_size=30, p=2, metric='minkowski'. SVM: C=1.0, kernel='rbf', degree=3, gamma='auto', shrinking=True, probability=False, class_weight=None. RF: n_estimators=100, criterion='gini', max_depth=None, min_samples_split=2, min_samples_leaf=1, min_weight_fraction_leaf=0.0, max_features='sqrt', bootstrap=True, class_weight=None. The parameters of the three classifiers are kept the same in all experiments. Therefore, the description of these parameters is omitted in the following sections.

To investigate the impact of the test set size on classification performance, we set the test set size to 0.2, 0.5, and 0.8. Table 1 presents the classification results accordingly. Table 1 shows that KNN is the best classifier among the three, achieving 99.41%, 99.35% and 99.07% accuracy on lead 1 for the three experiments with different test set sizes. The classification metrics of each classifier decrease slightly as the test set size grows (i.e. training set size decreases). For example, the F1-score using KNN is 0.9895 and 0.9837 at test set sizes of 0.2 and 0.8, respectively, a difference of only 0.58 percentage points. This clearly shows that heartbeat dynamics has a good ability to discriminate between heartbeats and achieves convincing classification results even with a training set to test set ratio of 2:8. Moreover, the classification performance on lead 1 performs better than that on lead 2. The main reason is that the second channel of the 48 ECG recordings was measured from multiple different leads (lead V1: 40 recordings; lead V2: 4 recordings; lead V4: 1 recording; lead V5: 2 recordings; modified lead II: 1 recording).

To show the classification details for each heartbeat class, we give the classification confusion matrix (Figure 2) and the classification report (Table 2) using KNN as the classifier at the test set size of 0.2. It can be seen that the classification performance for PVC is relatively poor. On lead 1, 50 PVCs are misclassified as normal heartbeats, with a recall of only 95.90%, about 3.5 percentage points lower than the recall for the other heartbeat classes. The main reason for this is that the patterns among the PVC heartbeats can significantly differ from each other [60], [61] makes the classification of PVC more challenging. The poor classification performance of PVC heartbeats affects the overall classification performance of this experiment.

Another 5-fold cross-validation experiment was conducted on the original heartbeat dataset with KNN as the classifier to avoid classification bias due to the choice of training heartbeats, and the results are shown in Figure 3. Figure 3 shows that the variation of all classification metrics across the fold experiments is small, indicating that the KNN classifier using heartbeat dynamics has good generalization ability. The mean values of accuracy, precision, recall and F1-score for the 5-fold cross-validation experiment on lead 1 (lead 2) are 99.41%, 99.10%, 98.84% and 0.9897 (99.11%, 98.61%, 98.23%, and 0.9841), respectively.

(a) Lead 1

(b) Lead 2

FIGURE 2. Confusion matrix for KNN classification on the original heartbeat dataset: Test set size of 0.2.

B. EXPERIMENTS ON BALANCED HEARTBEAT DATASET

The heartbeat dataset suffers from a class imbalance problem. It is a common problem in various fields (e.g., text classification, bioinformatics, medical diagnosis), which commonly affects the performance of both learning and classification algorithms [62]. Resampling is the most common approach to address class imbalance, either by oversampling the majority class or undersampling the minority class. To investigate the impact of class imbalance on heartbeat classification performance when using heartbeat dynamics as the exclusive classification feature, the edited nearest neighbours (ENN) method [63] and synthetic minority oversampling technique (SMOTE) [64] are performed on the extracted heartbeat dynamics to balance the distribution of heartbeat classes.

TABLE 1. Classifier performance on the original heartbeat dataset, with macro-averages for metrics (excluding accuracy) across all heartbeat classes.

	Lead 1											
Test set size	KNN				RF				SVM			
	Precision	Recall	Accuracy	F1-score	Precision	Recall	Accuracy	F1-score	Precision	Recall	Accuracy	F1-score
0.2	99.12%	98.78%	99.41%	0.9895	99.16%	97.48%	99.01%	0.9830	98.83%	97.44%	99.01%	0.9812
0.5	99.03%	98.68%	99.35%	0.9885	99.17%	97.25%	98.94%	0.9819	98.78%	97.24%	98.92%	0.9800
0.8	98.70%	98.06%	99.07%	0.9837	98.69%	95.86%	98.51%	0.9718	98.35%	96.46%	98.54%	0.9737
	Lead 2											
Test set size	KNN				RF				SVM			
	Precision	Recall	Accuracy	F1-score	Precision	Recall	Accuracy	F1-score	Precision	Recall	Accuracy	F1-score
0.2	98.51%	98.14%	99.08%	0.9831	98.95%	96.71%	98.75%	0.9776	97.90%	95.05%	98.10%	0.9629
0.5	98.43%	97.99%	99.01%	0.9820	98.69%	95.86%	98.42%	0.9718	97.63%	94.25%	97.82%	0.9572
0.8	98.09%	97.44%	98.77%	0.9773	98.61%	95.16%	98.18%	0.9674	97.31%	93.35%	97.51%	0.9502

TABLE 2. KNN classification report on the original heartbeat dataset: Test set size of 0.2, where "support" represents the total number of samples in the test set for the respective class.

	Lead 1				Lead 2			
	Precision	Recall	F1-score	Support	Precision	Recall	F1-score	Support
Normal	99.57%	99.74%	0.9965	14966	99.39%	99.57%	0.9948	14953
LBBB	98.91%	99.27%	0.9909	1643	97.26%	99.13%	0.9819	1612
RBBB	99.25%	99.25%	0.9925	1461	99.65%	99.86%	0.9976	1443
PVC	97.94%	95.90%	0.9691	1389	96.44%	92.40%	0.9438	1435
PB	99.93%	99.73%	0.9983	1462	99.80%	99.73%	0.9976	1478
Macro avg	99.12%	98.78%	0.9895	20921	98.51%	98.14%	0.9831	20921
Weighted avg	99.41%	99.41%	0.9941	20921	99.07%	99.08%	0.9907	20921
Accuracy	99.41%				99.08%			

TABLE 3. Distribution of different classes of heartbeats in the training set, undersampled training set and test set.

		Lead 1					
Test set size		Normal	LBBB	RBBB	PVC	PB	Total
0.2	Traing set	60050	6483	5740	5706	5701	83680
	Undersampled training set	59588	6406	5684	5322	5701	82701
	Test set	14970	1589	1515	1418	1429	20921
0.5	Traing set	37533	4031	3536	3549	3651	52300
	Undersampled training set	37206	3971	3536	3261	3642	51616
	Test set	37487	4041	3719	3575	3479	52301
0.8	Traing set	15016	1621	1455	1470	1358	20920
	Undersampled training set	14842	1589	1427	1325	1358	20541
	Test set	60004	6451	5800	5654	5772	83681
		Lead 2					
Test set size		Normal	LBBB	RBBB	PVC	PB	Total
0.2	Traing set	60047	6438	5780	5744	5671	83680
	Undersampled training set	59383	6328	5764	5080	5671	82226
	Test set	14973	1634	1475	1380	1459	20921
0.5	Traing set	37608	4051	3598	3474	3569	52300
	Undersampled training set	37164	3973	3587	3474	3550	51748
	Test set	37412	4021	3657	3650	3561	52301
0.8	Traing set	15046	1638	1407	1375	1454	20920
	Undersampled training set	14864	1597	1405	1375	1444	20685
	Test set	59974	6434	5848	5749	5676	83681

1) UNDERSAMPLING

ENN is an undersampling method that removes data that are different from most of the nearest neighbors (i.e., noisy data). In this subsection, we will use the ENN method on the extracted heartbeat dynamics of the training set to balance the distribution of the heartbeat classes. As in the experiments on the original heartbeat dataset, we set the size of the test set to 0.2, 0.5, and 0.8. Table 3 displays the distribution of different classes of heartbeats in the training set, undersampled training set, and test set.

Table 4 shows the classification performance of KNN, SVM, and RF classifiers with heartbeat dynamics as input features on the balanced heartbeat dataset after using the ENN method. Comparing Tables 1 and 4, we find that: (1) for the KNN classifier, the classification metrics for all experiments are lower than those on the original heartbeat data; (2) for the RF classifier, all classification metrics also decrease for all experiments except for the slight improvement in classification recall, accuracy, and F1-score on lead 2 with a test set size of 0.5; (3) for the SVM classifier, there is an improvement in recall for some experiments, but the overall classification performance also decreases. Moreover, the classification performance of KNN on the undersampled heartbeat dataset is still the best among the three classifiers.

TABLE 4. Classifier performance on the undersampled heartbeat dataset, with macro-averages for metrics (excluding accuracy) across all heartbeat classes.

Lead 1												
Test set size	KNN				RF				SVM			
	Precision	Recall	Accuracy	F1-score	Precision	Recall	Accuracy	F1-score	Precision	Recall	Accuracy	F1-score
0.2	98.91%	98.58%	99.28%	0.9873	98.96%	97.38%	98.91%	0.9815	98.57%	97.53%	98.95%	0.9804
0.5	98.80%	98.33%	99.19%	0.9855	98.99%	96.97%	98.79%	0.9795	98.45%	97.28%	98.84%	0.9785
0.8	98.22%	97.66%	98.84%	0.9793	98.31%	95.71%	98.25%	0.9697	97.92%	96.20%	98.37%	0.9703
Lead 2												
Test set size	KNN				SVM				RF			
	Precision	Recall	Accuracy	F1-score	Precision	Recall	Accuracy	F1-score	Precision	Recall	Accuracy	F1-score
0.2	98.50%	97.91%	99.01%	0.9819	98.66%	96.09%	98.52%	0.9729	97.90%	94.77%	98.04%	0.9616
0.5	98.04%	97.88%	98.86%	0.9795	98.49%	96.21%	98.47%	0.9729	97.51%	94.53%	97.83%	0.9583
0.8	97.47%	97.13%	98.53%	0.9729	98.07%	94.88%	98.01%	0.9636	97.79%	97.83%	97.83%	0.9774

TABLE 5. KNN classification report on the undersampled heartbeat dataset: Test set size of 0.2, where "support" represents the total number of samples in the test set for the respective class.

	Lead 1				Lead 2			
	Precision	Recall	F1-score	Support	Precision	Recall	F1-score	Support
Normal	99.49%	99.64%	0.9956	14970	99.27%	99.55%	0.9941	14973
LBBB	98.44%	99.37%	0.9890	1589	97.59%	99.08%	0.9833	1634
RBBB	98.56%	99.67%	0.9911	1515	99.53%	99.93%	0.9973	1475
PVC	98.53%	94.50%	0.9647	1418	96.48%	91.38%	0.9386	1380
PB	99.51%	99.72%	0.9962	1429	99.66%	99.59%	0.9962	1459
Macro avg	98.91%	98.58%	0.9873	20921	98.50%	97.91%	0.9819	20921
Weighted avg	99.28%	99.28%	0.9927	20921	99.00%	99.01%	0.9900	20921
Accuracy	99.28%				99.01%			

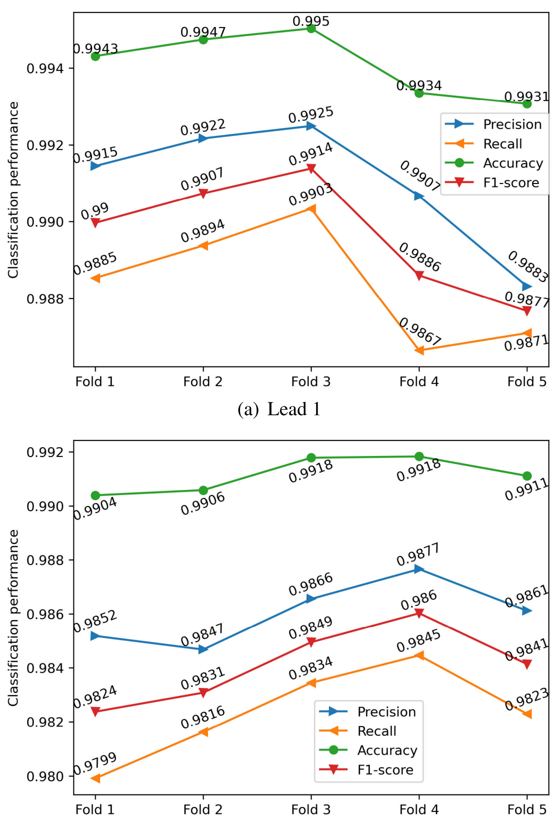


FIGURE 3. Classification results of a 5-fold cross-validation experiment: KNN on the original heartbeat dataset.

To understand the details of the change in classification performance, Table 5 and Figure 4 show the classification report and classification confusion matrix using KNN as the classifier on the undersampled heartbeat dataset with a test set size of 0.2. Comparing Tables 2 and 5, we can see a slight decrease in most of the classification metrics across all classes of heartbeat, making the overall classification performance on the undersampled heartbeat dataset lower than that on the original heartbeat dataset. Among them, the classification metrics for PVCs are somewhat special, with a slight increase in classification precision for PVCs on lead 1 and lead 2, but a larger decrease in recall compared to the increase in precision, resulting in a decrease in F1-score as well. We believe that this is mainly because heartbeat dynamics are more sensitive to noise, causing some heartbeats of one class to fall in clusters of other classes in the dynamics space. Undersampling the heartbeat dynamics of the training set may discard the training samples that are closest to some samples in the test set.

To avoid classification bias due to the choice of training heartbeats, we construct a 5-fold cross-validation experiment on the undersampled heartbeat dataset. The classification performance metrics under different folds are shown in Figure 5. It can be found that there are minor differences in the performance metrics for different folds. The mean values of accuracy, precision, recall and F1-score for the 5-fold cross-validation experiment on lead 1 (lead 2) are 99.35%, 98.92%, 98.85% and 0.9888 (99.04%, 98.29%, 98.29%, and 0.9829), respectively.

2) OVERSAMPLING

In this subsection, we attempt to address the class imbalance problem using an oversampling approach. SMOTE is employed to balance the heartbeat classes. SMOTE is an

TABLE 6. Distribution of different classes of heartbeats in the training set, oversampled training set and test set.

Lead 1							
Test set size		Normal	LBBB	RBBB	PVC	PB	Total
0.2	Training set	60057	6459	5778	5688	5698	83680
	Oversampled training set	59447	60049	60052	60020	60049	299617
	Test set	14963	1613	1477	1436	1432	20921
0.5	Training set	37549	4077	3627	3532	3515	52300
	Oversampled training set	37082	37544	37549	37528	37546	187249
	Test set	37471	3995	3628	3592	3615	52301
0.8	Training set	15032	1655	1456	1386	1391	20920
	Oversampled training set	14780	15027	15030	15024	15030	74891
	Test set	59988	6417	5799	5738	5739	83681
Lead 2							
Test set size		Normal	LBBB	RBBB	PVC	PB	Total
0.2	Training set	60033	6479	5774	5698	5696	83680
	Oversampled training set	59105	60012	60025	59940	60032	299114
	Test set	14987	1593	1481	1426	1434	20921
0.5	Training set	37412	4064	3674	3562	3588	52300
	Oversampled training set	36814	37393	37399	37360	37411	186377
	Test set	37608	4008	3581	3562	3542	52301
0.8	Training set	15057	1564	1433	1423	1443	20920
	Oversampled training set	14764	15057	15056	15028	15054	74959
	Test set	59963	6508	5822	5701	5687	83681

TABLE 7. Classifier performance on the oversampled heartbeat dataset, with macro-averages for metrics (excluding accuracy) across all heartbeat classes.

Lead 1												
Test set size	KNN				RF				SVM			
	Precision	Recall	Accuracy	F1-score	Precision	Recall	Accuracy	F1-score	Precision	Recall	Accuracy	F1-score
0.2	98.06%	99.25%	99.19%	0.9864	98.32%	98.75%	99.14%	0.9852	97.16%	99.05%	98.82%	0.9807
0.5	97.91%	99.09%	99.10%	0.9849	98.24%	98.60%	99.05%	0.9841	96.34%	98.74%	98.49%	0.9748
0.8	97.07%	98.86%	98.77%	0.9795	97.73%	98.03%	98.75%	0.9787	94.35%	98.07%	97.60%	0.9608
Lead 2												
Test set size	KNN				RF				SVM			
	Precision	Recall	Accuracy	F1-score	Precision	Recall	Accuracy	F1-score	Precision	Recall	Accuracy	F1-score
0.2	97.08%	98.70%	98.76%	0.9786	97.92%	98.18%	98.89%	0.9805	94.00%	97.75%	97.17%	0.9564
0.5	96.86%	98.43%	98.61%	0.9762	97.85%	97.90%	98.79%	0.9787	94.85%	97.43%	97.58%	0.9605
0.8	96.31%	98.30%	98.40%	0.9727	97.57%	97.10%	98.49%	0.9733	95.56%	96.74%	97.77%	0.9614

oversampling strategy that combines existing and created synthetic samples. However, some noisy data can be generated when oversampling with the SMOTE method. Thus, the oversampled data needs to be cleaned. Here we choose the ENN method to remove the noisy data generated during oversampling. As in the experiments on the original and undersampled heartbeat datasets, we set the size of the test set to 0.2, 0.5, and 0.8. Table 6 displays the distribution of different classes of heartbeats in the training set, oversampled training set, and test set.

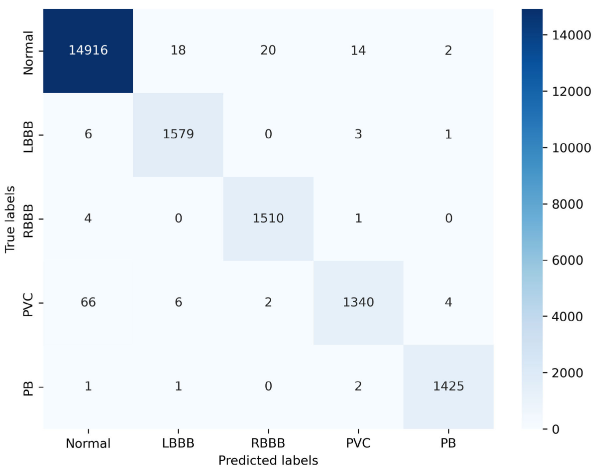
Table 7 presents the classification results for KNN, SVM and RF with different test set sizes on the oversampled heartbeat dataset. Comparing Tables 1 and 7, it can be seen that recall increases for all experiments on the oversampled heartbeat dataset, but precision decreases more, resulting in a decrease in both accuracy and F1-score. In addition, it can be found in comparing Tables 4 and 7 that the overall classification performance on the oversampled heartbeat dataset is lower than that on the undersampled heartbeat dataset. Furthermore, the overall classification performance on the oversampled heartbeat dataset is lower than that on the undersampled and original heartbeat datasets.

Specially, to understand the details of the classification performance change, we show the classification confusion matrix (Figure 6) and the classification report (Table 8) with KNN as the classifier at the test set size of 0.2. Comparing Tables 2 and 8, we can see that on the oversampled heartbeat dataset, the classification precision increases but the recall decreases for normal heartbeats, while the opposite is true for the remaining four minority classes of heartbeats, where the classification recall increases but the precision decreases. In particular, for PVCs, there is a large improvement in classification recall, but also a large decrease in classification accuracy. Taking the classification performance on lead 1 as an example, the classification precision for PVC on the original and oversampled heartbeat dataset is 97.94% and 95.20%, respectively, and the classification recall is 95.90% and 97.98%, respectively. We believe that this is because the negative impact of the noisy data generated during the oversampling process outweighs the positive impact of the balanced treatment (using oversampling method) on the classification performance.

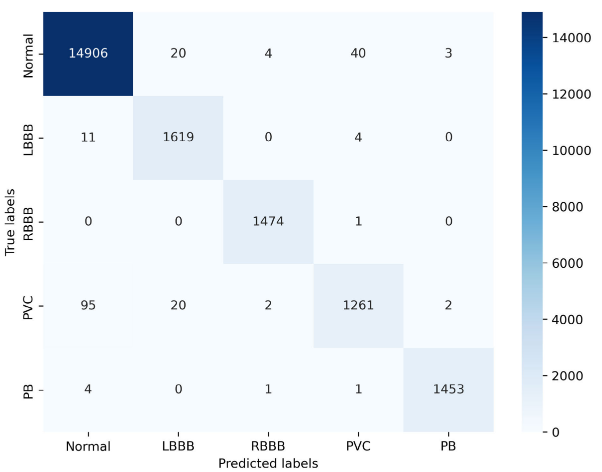
To avoid classification bias due to the choice of training heartbeats, we conduct a 5-fold cross-validation experiment to further evaluate the classification performance using

TABLE 8. KNN classification report on the oversampled heartbeat dataset: Test set size of 0.2, where "support" represents the total number of samples in the test set for the respective class.

	Lead 1				Lead 2			
	Precision	Recall	F1-score	Support	Precision	Recall	F1-score	Support
Normal	99.84%	99.16%	0.9950	14963	99.74%	98.78%	0.9926	14987
LBBB	97.93%	99.50%	0.9871	1613	97.47%	99.00%	0.9822	1593
RBBB	97.87%	99.73%	0.9879	1477	99.19%	99.80%	0.9950	1481
PVC	95.20%	97.98%	0.9657	1436	89.69%	96.42%	0.9294	1426
PB	99.44%	99.86%	0.9965	1432	99.30%	99.51%	0.9941	1434
Macro avg	98.06%	99.25%	0.9864	20921	97.08%	98.70%	0.9786	20921
Weighted avg	99.21%	99.19%	0.9920	20921	98.81%	98.76%	0.9877	20921
Accuracy	99.19%				98.76%			

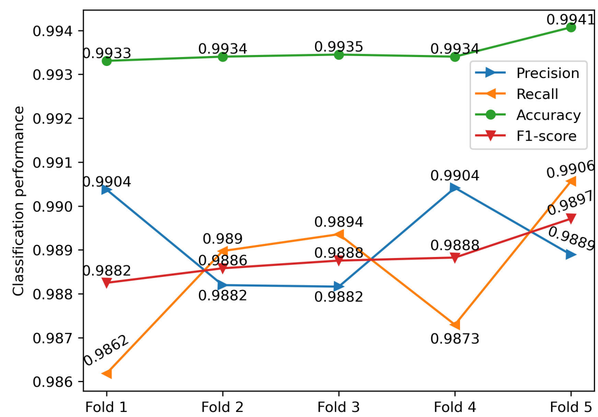


(a) Lead 1

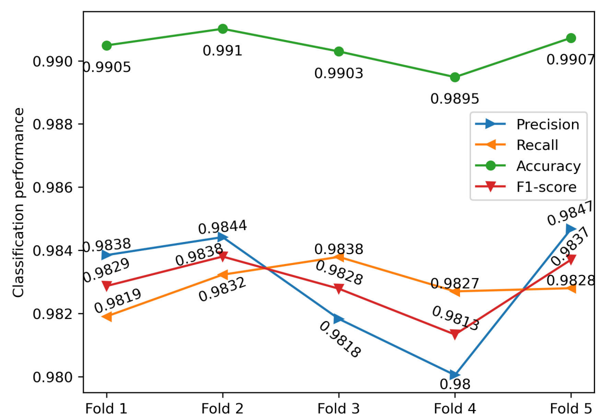


(b) Lead 2

FIGURE 4. Confusion matrix for KNN classification on the undersampled heartbeat dataset: Test set size of 0.2.



(a) Lead 1

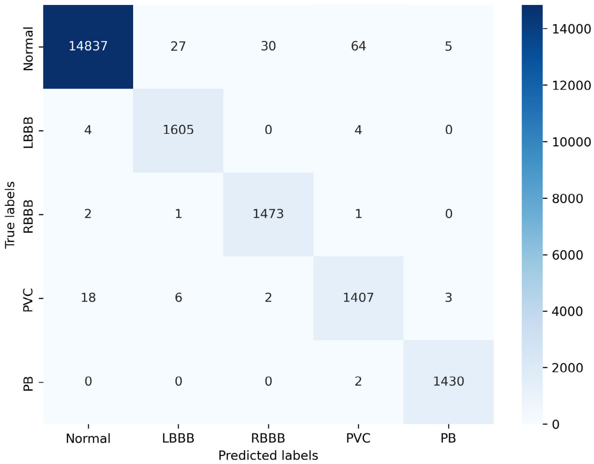


(b) Lead 2

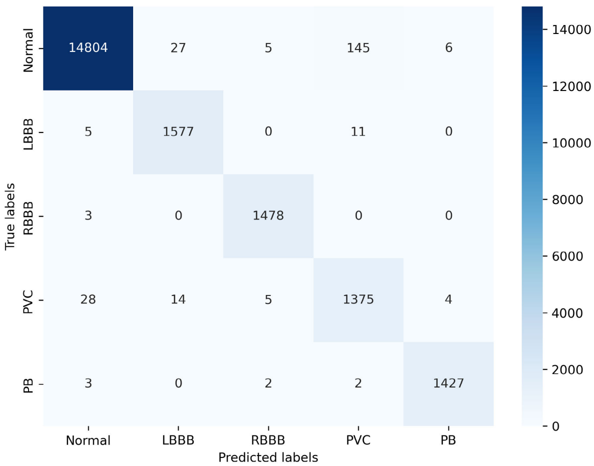
FIGURE 5. Classification results of a 5-fold cross-validation experiment: KNN on the undersampled heartbeat dataset.

heartbeat dynamics as input features on an oversampled heartbeat dataset. Figure 7 demonstrates the classification results of the 5-fold cross-validation experiment. It can be seen that all metrics have the same trend of variation across different fold experiments. The mean values of accuracy, precision, recall and F1-score for the 5-fold cross-validation experiment on lead 1 (lead 2) are 99.22%, 98.12%, 99.29% and 0.9870 (98.76%, 97.16%, 98.71%, 0.9791), respectively.

A comprehensive analysis of all experiments reveals that: (1) For KNN classifier: Undersampling makes all classification metrics lower. Oversampling increases recall and decreases accuracy, precision, and F1-score. The overall classification performance is highest on the original heartbeat dataset and lowest on the oversampled heartbeat dataset; (2) For RF classifier: Undersampling reduces all classification metrics. Oversampling reduces classification precision, but improves recall considerably, resulting in higher accuracy and F1-score. The overall classification performance is

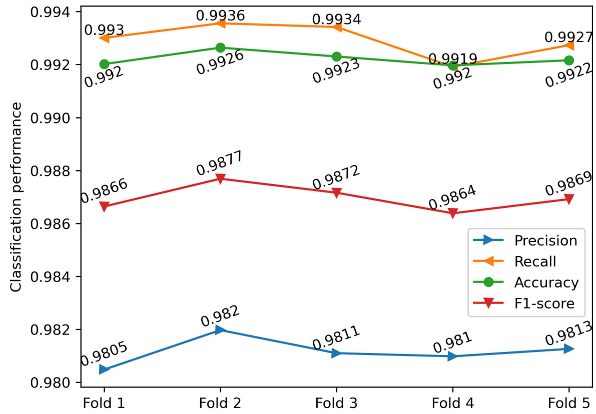


(a) Lead 1

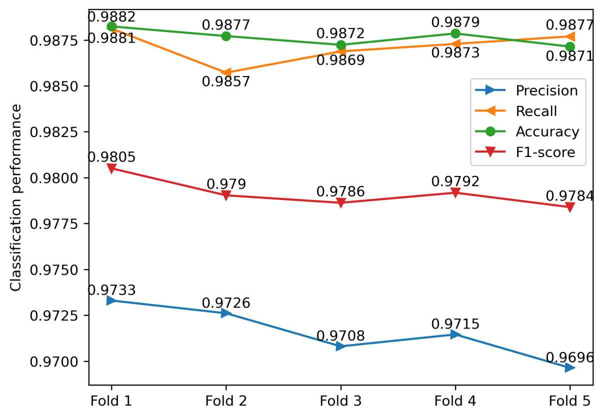


(b) Lead 2

FIGURE 6. **Confusion matrix for KNN classification on the oversampled heartbeat dataset: Test set size of 0.2.**



(a) Lead 1



(b) Lead 2

FIGURE 7. **Classification results of a 5-fold cross-validation experiment: KNN on the oversampled heartbeat dataset.**

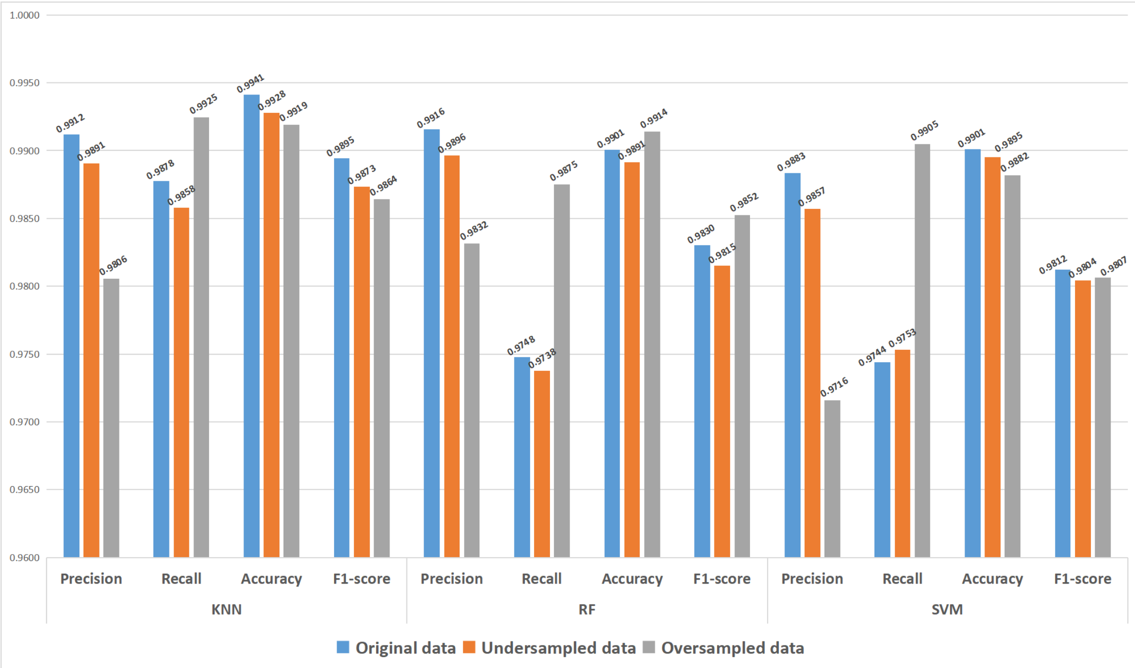
highest on the oversampled heartbeat dataset and lowest on the undersampled heartbeat dataset; (3) For the SVM classifier: Undersampling improves recall, but precision, accuracy and F1-score all decrease. Oversampling considerably reduces the precision and also highly improves the recall, with a slight decrease in accuracy and F1-score. The overall classification performance is highest on the original heartbeat dataset and lowest on the oversampled heartbeat dataset; (4) Using heartbeat dynamics as the classification feature, KNN has the best classification performance among these three classifiers. Figure 8 depicts classification metrics for KNN, SVM, and RF with test set size 0.2 to visualize classification performance based on heartbeat dynamics on the original, undersampled, and oversampled heartbeat datasets.

IV. DISCUSSION

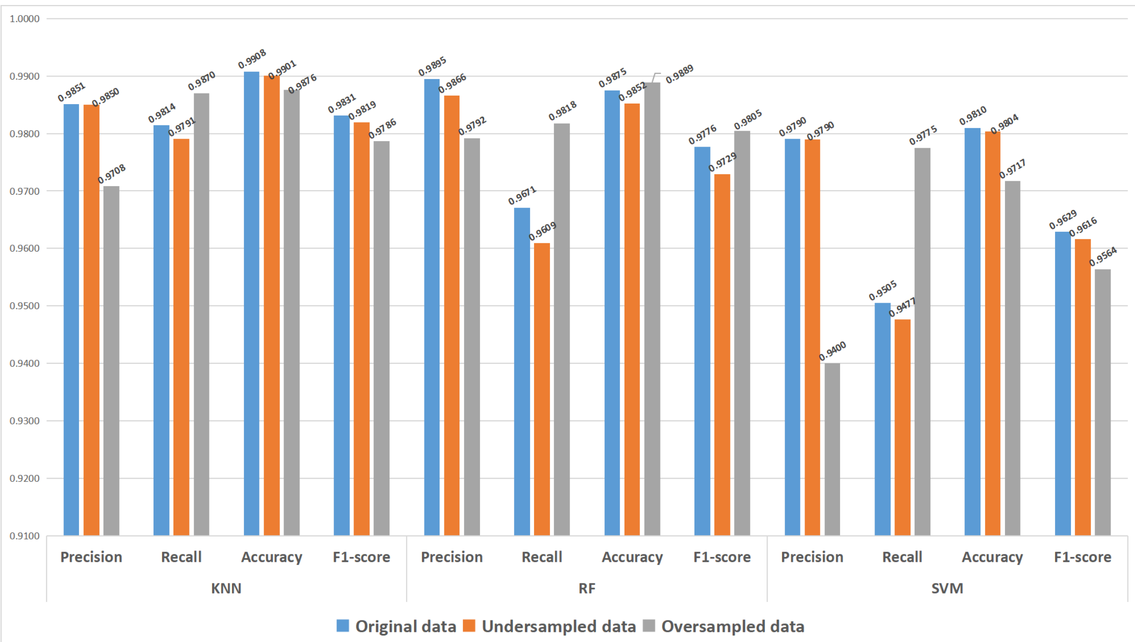
Experimental results show that using heartbeat dynamics as the exclusive classification feature achieves outstanding performance for heartbeat classification. Although the original heartbeat dataset suffers from data class imbalance issue, a simple classifier KNN also achieves promising classification performance with average accuracy, precision, recall and F1-score of 99.41%, 99.10%, 98.84% and 0.9899. To investigate the impact of class imbalance on classification performance using heartbeat dynamics as classification features, two methods, ENN and SMOTE, were selected to balance the original heartbeat dataset.

While, experimental results show that the balanced treatment of heartbeat dataset does not improve the performance of heartbeat classification. This may be due to the fact that some important data is lost during undersampling, leading to underfitting. In the oversampling process, a large number of minority class heartbeats are synthesized, which also contains some noisy data, resulting in misclassification of heartbeats at the classification boundary. In other words, the effect of noisy data on classification performance outweighs the effect of class imbalance. From another perspective, heartbeat dynamics is highly discriminative and algorithms employing it as a classification feature are insensitive to the class imbalance problem.

We directly compare the proposed method with several state-of-the-art heartbeat classification methods to demonstrate its advantage. Table 9 summarizes the experimental



(a) Lead 1



(b) Lead 2

FIGURE 8. **Comparison of classification performance on the original, undersampled and oversampled heartbeat datasets: Test set size of 0.2.**

results of several state-of-the-art methods for heartbeat classification using the MIT-BIH arrhythmia database. Compared with other studies, the proposed method has the same level of accuracy as the references [65] and [66]. The precision and recall of the classification is only lower than that of the reference [65]. While our method's classification performance lags behind that reported in the literature [67], it is important to note that our method utilizes only single-lead ECG signals, which require comparatively less computational effort and render it better suited for deployment on wearable devices. In addition, the proposed method does not perform any tuning operation on the parameters of the classifier. It confirms that heartbeat dynamics is a salient feature for heartbeat classification, and heartbeat classification algorithms

TABLE 9. Performance comparison of the proposed method with state-of-the-art studies.

Literature	Year	# of Classes	Techniques	Accuracy	Recall	Precision
[37]	2023	4	EMD+CNN	99.01%	99.11%	99.01%
[40]	2020	5	SpEn+ 2D^{2}PCA+DCNN	98.33%	98.33%	98.34%
[68]	2018	5	CNN+LSTM	98.1%	97.5%	
[69]	2019	5	M-U-net	97.32%	94.44%	94.7%
[65]	2022	7	DeepNet	99.56%	99.26%	99.42%
[66]	2022	5	TFDF+1D-CNN	99.43%		
[70]	2019	5	HOS+DL	97.8%	97.8%	97.8%
[71]	2022	5	LSTM+AE	98.57%	97.89%	97.55%
[67]	2023	5	SWT+Bi-LSTM	99.72%	99.02%	99.22%
Proposed		5	DetL+PCA+KNN	99.41%	98.84%	99.10%

Abbreviations of features and techniques: SpEn: spectral entropy; 2D^{2}PCA: two-directional two-dimensional principal component analysis; DCNN: deep CNN; M-U-net: modified U-net; TFDF: time-frequency domain fusion; 1D-CNN: 1-dimensional CNN; HOS: higher order spectral; DL: deep learning; LSTM: long short-term memory networks; AE: autoencoder; DetL: deterministic learning; SWT: stationary wavelet transform; Bi-LSTM: bi-directional LSTM.

based on heartbeat dynamics will have better generalization capabilities.

Compared with existing related studies, this study has the following advantages or features: (1) heartbeat dynamics is a new method to describe ECG more comprehensively and is used as an exclusive classification feature; (2) it overcomes the difficulties of static feature combination and optimization in traditional machine learning approaches; (3) it is interpretable compared with DL-extracted features; (4) heartbeat dynamics has good discriminating for different classes of heartbeats, making it easy to deploy into wearable devices with simple classifiers; and (5) it is insensitive to the case of class imbalance.

However, since ECG is susceptible to various noise contamination, even with various denoising methods, it is not guaranteed to obtain a completely clean heartbeat signal. While the heartbeat dynamics accurately reflect the dynamical changes in the ECG, it is more sensitive to various anomalies in the depolarization and repolarization processes of the cardiocytes. Thus, it is also more sensitive to noise, which can affect the classification performance of heartbeats. In the next research, obtaining purer ECG signals and weakening the effect of noise on heartbeat dynamics are important directions for our efforts.

V. CONCLUSION

In this paper, we propose a different heartbeat classification feature called heartbeat dynamics. It provides a comprehensive characterization of the heartbeat in a new way that is more sensitive to weak changes in the heartbeat. We evaluate the ability of heartbeat dynamics to classify heartbeats using three classical, simple, and traditional classifiers, KNN, SVM, and RF. We achieved 99.41% accuracy, 99.10% precision, 98.84% recall, and 0.9897 F1-score with KNN as classifier on the original heartbeat dataset. Experimental results on both original and balanced datasets show that the heartbeat dynamics has a strong ability to discriminate between different classes of heartbeats and is insensitive to whether the data is balanced or not. We anticipate that the heartbeat dynamics feature will enhance the generalization capacity of the arrhythmia detection algorithm when integrated with other static features.

As mentioned in the introduction section, the classification performance of DL models has fundamentally reached its limit and it is necessary to investigate feature extraction from a different perspective. Heartbeat dynamics provides an adequate characterization of heartbeats from a different perspective. In the next study, we intend to combine the heartbeat dynamics with the DL method and fuse it with the features extracted from the heartbeat by the DL model. This is expected to further improve the performance of arrhythmia classification.

REFERENCES

[1] M. Liyuan, W. Zengwu, F. Jing, and H. Shengshou, "An essential introduction to the annual report on cardiovascular health and diseases in China (2021)," *Chin. Gen. Pract.*, vol. 25, no. 27, p. 3331, 2022.

[2] I. Hussain and S. J. Park, "Big-ECG: Cardiographic predictive cyber-physical system for stroke management," *IEEE Access*, vol. 9, pp. 123146–123164, 2021.

[3] M. S. Islam, I. Hussain, M. M. Rahman, S. J. Park, and M. A. Hossain, "Explainable artificial intelligence model for stroke prediction using EEG signal," *Sensors*, vol. 22, no. 24, p. 9859, Dec. 2022.

[4] J. Zhu, Y. Zeng, J. Zhou, and X. Dong, "ECG heartbeat classification based on combined features extracted by PCA, KPCA, AKPCA and DWT," in *Proc. IEEE 35th Int. Symp. Comput.-Based Med. Syst. (CBMS)*, 2022, pp. 155–159.

[5] R. J. Martis, U. R. Acharya, K. M. Mandana, A. K. Ray, and C. Chakraborty, "Application of principal component analysis to ECG signals for automated diagnosis of cardiac health," *Exp. Syst. Appl.*, vol. 39, no. 14, pp. 11792–11800, Oct. 2012.

[6] R. J. Martis, U. R. Acharya, C. M. Lim, K. M. Mandana, A. K. Ray, and C. Chakraborty, "Application of higher order cumulant features for cardiac health diagnosis using ECG signals," *Int. J. Neural Syst.*, vol. 23, no. 4, Aug. 2013, Art. no. 1350014.

[7] A. Daamouche, L. Hamami, N. Alajlan, and F. Melgani, "A wavelet optimization approach for ECG signal classification," *Biomed. Signal Process. Control*, vol. 7, no. 4, pp. 342–349, Jul. 2012.

[8] U. Satija, B. Ramkumar, and M. S. Manikandan, "An automated ECG signal quality assessment method for unsupervised diagnostic systems," *Biocybernetics Biomed. Eng.*, vol. 38, no. 1, pp. 54–70, 2018.

[9] M. Abdelazez, S. Rajan, and A. D. C. Chan, "Detection of atrial fibrillation in compressively sensed electrocardiogram measurements," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–9, 2021.

[10] H. M. Rai, A. Trivedi, and S. Shukla, "ECG signal processing for abnormalities detection using multi-resolution wavelet transform and artificial neural network classifier," *Measurement*, vol. 46, no. 9, pp. 3238–3246, Nov. 2013.

[11] N. Sinha and A. Das, "Automatic diagnosis of cardiac arrhythmias based on three stage feature fusion and classification model using DWT," *Biomed. Signal Process. Control*, vol. 62, Sep. 2020, Art. no. 102066.

[12] E. Alickovic and A. Subasi, "Medical decision support system for diagnosis of heart arrhythmia using DWT and random forests classifier," *J. Med. Syst.*, vol. 40, no. 4, pp. 1–12, Apr. 2016.

[13] H. Li, D. Yuan, Y. Wang, D. Cui, and L. Cao, "Arrhythmia classification based on multi-domain feature extraction for an ECG recognition system," *Sensors*, vol. 16, no. 10, p. 1744, Oct. 2016.

[14] M. Ramkumar, C. Ganesh Babu, K. V. Kumar, D. Hepsiba, A. Manjunathan, and R. S. Kumar, "ECG cardiac arrhythmias classification using DWT, ICA and MLP neural networks," *J. Phys., Conf.*, vol. 1831, no. 1, Mar. 2021, Art. no. 012015.

[15] T. Li and M. Zhou, "ECG classification using wavelet packet entropy and random forests," *Entropy*, vol. 18, no. 8, p. 285, Aug. 2016.

[16] F. A. Elhaj, N. Salim, A. R. Harris, T. T. Swee, and T. Ahmed, "Arrhythmia recognition and classification using combined linear and nonlinear features of ECG signals," *Comput. Methods Programs Biomed.*, vol. 127, pp. 52–63, Apr. 2016.

[17] C. Kuang Chua, "Cardiac health diagnosis using higher order spectra and support vector machine," *Open Med. Informat. J.*, vol. 3, no. 1, pp. 1–8, Feb. 2009.

[18] R. J. Martis, U. R. Acharya, K. M. Mandana, A. K. Ray, and C. Chakraborty, "Cardiac decision making using higher order spectra," *Biomed. Signal Process. Control*, vol. 8, no. 2, pp. 193–203, Mar. 2013.

[19] K. C. Chua, V. Chandran, U. R. Acharya, and C. M. Lim, "Cardiac state diagnosis using higher order spectra of heart rate variability," *J. Med. Eng. Technol.*, vol. 32, no. 2, pp. 145–155, Jan. 2008.

[20] S. Karimifard and A. Ahmadian, "A robust method for diagnosis of morphological arrhythmias based on Hermitian model of higher-order statistics," *Biomed. Eng. online*, vol. 10, no. 1, pp. 1–18, 2011.

[21] S. Osowski, L. T. Hoai, and T. Markiewicz, "Support vector machine-based expert system for reliable heartbeat recognition," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 4, pp. 582–589, Apr. 2004.

[22] G. de Lannoy, D. Francois, J. Delbeke, and M. Verleysen, "Weighted conditional random fields for supervised interpatient heartbeat classification," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 1, pp. 241–247, Jan. 2012.

[23] S. Banerjee and M. Mitra, "Application of cross wavelet transform for ECG pattern analysis and classification," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 2, pp. 326–333, Feb. 2014.

[24] A. Ebrahimzadeh, B. Shakiba, and A. Khazaee, "Detection of electrocardiogram signals using an efficient method," *Appl. Soft Comput.*, vol. 22, pp. 108–117, Sep. 2014.

[25] S. Raj and K. C. Ray, "Sparse representation of ECG signals for automated recognition of cardiac arrhythmias," *Exp. Syst. Appl.*, vol. 105, pp. 49–64, Sep. 2018.

[26] P. de Chazal, M. O'Dwyer, and R. B. Reilly, "Automatic classification of heartbeats using ECG morphology and heartbeat interval features," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 7, pp. 1196–1206, Jul. 2004.

[27] S. Sahoo, B. Kanungo, S. Behera, and S. Sabut, "Multiresolution wavelet transform based feature extraction and ECG classification to detect cardiac abnormalities," *Measurement*, vol. 108, pp. 55–66, Oct. 2017.

[28] S. Raj and K. C. Ray, "ECG signal analysis using DCT-based DOST and PSO optimized SVM," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 3, pp. 470–478, Mar. 2017.

[29] 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," *Biomed. Signal Process. Control*, vol. 47, pp. 41–48, Jan. 2019.

[30] G. G. N. Geweid and J. D. Z. Chen, "Automatic classification of atrial fibrillation from short single-lead ECG recordings using a hybrid approach of dual support vector machine," *Exp. Syst. Appl.*, vol. 198, Jul. 2022, Art. no. 116848.

[31] M. Badr, S. Al-Otaibi, N. Alturki, and T. Abir, "Detection of heart arrhythmia on electrocardiogram using artificial neural networks," *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–10, Aug. 2022.

[32] A. Y. Hannun, P. Rajpurkar, M. Haghpanahi, G. H. Tison, C. Bourn, M. P. Turakhia, and A. Y. Ng, "Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network," *Nature Med.*, vol. 25, no. 1, pp. 65–69, Jan. 2019.

[33] Ö. Yildirim, P. Plawiak, R.-S. Tan, and U. R. Acharya, "Arrhythmia detection using deep convolutional neural network with long duration ECG signals," *Comput. Biol. Med.*, vol. 102, pp. 411–420, Nov. 2018.

[34] M. Ashfaq Khan and Y. Kim, "Cardiac arrhythmia disease classification using LSTM deep learning approach," *Comput., Mater. Continua*, vol. 67, no. 1, pp. 427–443, 2021.

[35] H. Wang, H. Shi, X. Chen, L. Zhao, Y. Huang, and C. Liu, "An improved convolutional neural network based approach for automated heartbeat classification," *J. Med. Syst.*, vol. 44, no. 2, pp. 1–9, Feb. 2020.

[36] J. Cui, L. Wang, X. He, V. H. C. De Albuquerque, S. A. AlQahtani, and M. M. Hassan, "Deep learning-based multidimensional feature fusion for classification of ECG arrhythmia," *Neural Comput. Appl.*, vol. 35, pp. 16073–16087, Sep. 2021.

[37] Y. Li, J.-H. Luo, Q.-Y. Dai, J. K. Eshraghian, B. W.-K. Ling, C.-Y. Zheng, and X.-L. Wang, "A deep learning approach to cardiovascular disease classification using empirical mode decomposition for ECG feature extraction," *Biomed. Signal Process. Control*, vol. 79, Jan. 2023, Art. no. 104188.

[38] 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 Trans. Instrum. Meas.*, vol. 70, pp. 1–9, 2021.

[39] S. Singh, S. K. Pandey, U. Pawar, and R. R. Janghel, "Classification of ECG arrhythmia using recurrent neural networks," *Proc. Comput. Sci.*, vol. 132, pp. 1290–1297, Jan. 2018.

[40] A. Asgharzadeh-Bonab, M. C. Amirani, and A. Mehri, "Spectral entropy and deep convolutional neural network for ECG beat classification," *Biocybernetics Biomed. Eng.*, vol. 40, no. 2, pp. 691–700, Apr. 2020.

[41] Y. Li, R. Qian, and K. Li, "Inter-patient arrhythmia classification with improved deep residual convolutional neural network," *Comput. Methods Programs Biomed.*, vol. 214, Feb. 2022, Art. no. 106582.

[42] J.-K. Kim, S. Jung, J. Park, and S. W. Han, "Arrhythmia detection model using modified DenseNet for comprehensible grad-CAM visualization," *Biomed. Signal Process. Control*, vol. 73, Mar. 2022, Art. no. 103408.

[43] Y. K. Kim, M. Lee, H. S. Song, and S.-W. Lee, "Automatic cardiac arrhythmia classification using residual network combined with long short-term memory," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–17, 2022.

[44] P. Xiong, S. M.-Y. Lee, and G. Chan, "Deep learning for detecting and locating myocardial infarction by electrocardiogram: A literature review," *Frontiers Cardiovascular Med.*, vol. 9, Mar. 2022, Art. no. 860032.

[45] C. Wang, X. Dong, S. Ou, W. Wang, J. Hu, and F. Yang, "A new method for early detection of myocardial ischemia: Cardiodynamicsgram (CDG)," *Sci. China Inf. Sci.*, vol. 59, no. 1, pp. 1–11, Jan. 2016.

[46] M. Deng, M. Tang, C. Wang, L. Shan, L. Zhang, J. Zhang, W. Wu, and L. Xia, "Cardiodynamicsgram as a new diagnostic tool in coronary artery disease patients with nondiagnostic electrocardiograms," *Amer. J. Cardiol.*, vol. 119, no. 5, pp. 698–704, Mar. 2017.

[47] Y. Wang, J. Sun, K. Sun, L. Li, X. Yu, C. Wang, H. Gu, Q. Sun, and X. Wang, "ECG-based cardiodynamicsgram can reflect anomalous functional information in coronary artery disease," *Clin. Cardiol.*, vol. 46, no. 6, pp. 639–647, Jun. 2023.

[48] X. Dong, J. Zhang, H. Lai, M. Tang, S. Ou, J. Dou, and C. Wang, "The predictive value of cardiodynamicsgram in myocardial perfusion abnormalities," *PLoS ONE*, vol. 13, no. 12, Dec. 2018, Art. no. e0208859.

[49] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. 215–220, Jun. 2000.

[50] C. Wang and D. J. Hill, "Deterministic learning and rapid dynamical pattern recognition," *IEEE Trans. Neural Netw.*, vol. 18, no. 3, pp. 617–630, May 2007.

[51] C. Wang and D. J. Hill, "Learning from neural control," *IEEE Trans. Neural Netw.*, vol. 17, no. 1, pp. 130–146, Jan. 2006.

[52] C. Wang and D. J. Hill, *Deterministic Learning Theory: For Identiflcation, Recognition, and Conirol*. Boca Raton, FL, USA: CRC Press, 2009.

[53] Z. Zhang, J. Dong, X. Luo, K.-S. Choi, and X. Wu, "Heartbeat classification using disease-specific feature selection," *Comput. Biol. Med.*, vol. 46, pp. 79–89, Mar. 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S001048251300348X

[54] M. Ringnér, "What is principal component analysis?" *Nature Biotechnol.*, vol. 26, no. 3, pp. 303–304, Mar. 2008.

[55] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.

[56] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[57] L. Breiman, "Bagging predictors," *Mach. Learn.*, vol. 24, no. 2, pp. 123–140, Aug. 1996.
[58] P. Yang, D. Wang, W.-B. Zhao, L.-H. Fu, J.-L. Du, and H. Su, "Ensemble of kernel extreme learning machine based random forest classifiers for automatic heartbeat classification," *Biomed. Signal Process. Control*, vol. 63, Jan. 2021, Art. no. 102138.
[59] M. Sheykhmousa, M. Mahdianpari, H. Ghanbari, F. Mohammadimanesh, P. Ghamisi, and S. Homayouni, "Support vector machine versus random forest for remote sensing image classification: A meta-analysis and systematic review," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 6308–6325, 2020.
[60] Y. Jung and H. Kim, "Detection of PVC by using a wavelet-based statistical ECG monitoring procedure," *Biomed. Signal Process. Control*, vol. 36, pp. 176–182, Jul. 2017.
[61] M. L. Talbi and A. Charef, "PVC discrimination using the QRS power spectrum and self-organizing maps," *Comput. Methods Programs Biomed.*, vol. 94, no. 3, pp. 223–231, Jun. 2009.
[62] K. N. V. P. S. Rajesh and R. Dhuli, "Classification of imbalanced ECG beats using re-sampling techniques and AdaBoost ensemble classifier," *Biomed. Signal Process. Control*, vol. 41, pp. 242–254, Mar. 2018.
[63] D. L. Wilson, "Asymptotic properties of nearest neighbor rules using edited data," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-2, no. 3, pp. 408–421, Jul. 1972.
[64] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, Jun. 2002.
[65] Y. P. Sai and L. V. R. Kumari, "Cognitive assistant DeepNet model for detection of cardiac arrhythmia," *Biomed. Signal Process. Control*, vol. 71, Jan. 2022, Art. no. 103221.
[66] B. Wang, G. Chen, L. Rong, Y. Liu, A. Yu, X. He, T. Wen, Y. Zhang, and B. Hu, "Arrhythmia disease diagnosis based on ECG time-frequency domain fusion and convolutional neural network," *IEEE J. Transl. Eng. Health Med.*, vol. 11, pp. 116–125, 2023.
[67] L. D. Sharma, J. Rahul, A. Aggarwal, and V. K. Bohat, "An improved cardiac arrhythmia classification using stationary wavelet transform decomposed short duration QRS segment and bi-LSTM network," *Multidimensional Syst. Signal Process.*, vol. 34, no. 2, pp. 503–520, Jun. 2023.
[68] S. L. Oh, E. Y. K. Ng, R. S. Tan, and U. R. Acharya, "Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length heart beats," *Comput. Biol. Med.*, vol. 102, pp. 278–287, Nov. 2018.
[69] S. L. Oh, E. Y. K. Ng, R. S. Tan, and U. R. Acharya, "Automated beat-wise arrhythmia diagnosis using modified U-Net on extended electrocardiographic recordings with heterogeneous arrhythmia types," *Comput. Biol. Med.*, vol. 105, pp. 92–101, Feb. 2019.
[70] H. Alquran, A. M. Alqudah, I. Abu-Qasmieh, A. Al-Badarneh, and S. Almashaqbeh, "ECG classification using higher order spectral estimation and deep learning techniques," *Neural Netw. World*, vol. 29, no. 4, pp. 207–219, 2019.
[71] P. Liu, X. Sun, Y. Han, Z. He, W. Zhang, and C. Wu, "Arrhythmia classification of LSTM autoencoder based on time series anomaly detection," *Biomed. Signal Process. Control*, vol. 71, Jan. 2022, Art. no. 103228.

XUNDE DONG received the M.S. degree in mathematical and applied mathematical and the Ph.D. degree in control theory and control engineering from the South China University of Technology, Guangzhou, China, in 2010 and 2014, respectively. He is currently an Associate Professor with the School of Automation Science and Engineering, South China University of Technology. His research interests include biomedical signal processing, machine learning, deep learning, dynamical pattern recognition, and deterministic learning.

WENJIE SI received the B.S. and M.S. degrees in control theory and control engineering from Zhengzhou University, Zhengzhou, China, in 2008 and 2011, respectively, and the Ph.D. degree in control theory and control engineering from the South China University of Technology, Guangzhou, China, in 2015. He is currently an Associate Professor with the School of Electrical and Control Engineering, Henan University of Urban Construction, Pingdingshan, China. His current research interests include adaptive neural control, nonlinear control, and deterministic learning theory.