Mathematical Biosciences and Engineering

MBE, 20(7): 13061−13085.
DOI: 10.3934/mbe.2023582
Received: 29 January 2023
Revised: 20 May 2023
Accepted: 22 May 2023
Published: 05 June 2023

http://www.aimspress.com/journal/MBE

Research article

Entropy-based reliable non-invasive detection of coronary microvascular dysfunction using machine learning algorithm

Xiaoye Zhao[1,2,3,†], Yinlan Gong[4,†], Lihua Xu[5], Ling Xia[6,7], Jucheng Zhang[8], Dingchang Zheng[9], Zongbi Yao[10], Xinjie Zhang[10], Haicheng Wei[2], Jun Jiang[11], Haipeng Liu[9,*] and Jiandong Mao[1,2,3,*]

[1] School of Instrument Science and Opto-electronic Engineering, Hefei University of Technology, Hefei 230009, Anhui, China
[2] School of Electrical and Information Engineering, North Minzu University, Yinchuan 750001, Ningxia, China
[3] Key Laboratory of Atmospheric Environment Remote Sensing of Ningxia, Yinchuan 750001, Ningxia, China
[4] Institute of Wenzhou, Zhejiang University, Wenzhou 325000, Zhejiang, China
[5] Hangzhou Linghua Biotech Ltd, Hangzhou 310009, Zhejiang, China
[6] Key Laboratory for Biomedical Engineering of Ministry of Education, Hangzhou 310009, Zhejiang, China
[7] Institute of Biomedical Engineering, Zhejiang University, Hangzhou 310009, Zhejiang, China
[8] Department of Clinical Engineering, The Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310009, Zhejiang, China
[9] Research Centre for Intelligent Healthcare, Coventry University, Coventry, CV1 5FB, United Kingdom
[10] Department of Cardiology, Ningxia Hui Autonomous Region People's Hospital, Yinchuan 750021, Ningxia, China
[11] Department of Cardiology, The Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310009, Zhejiang, China

* **Correspondence:** Email: haipeng.liu@coventry.ac.uk, mao_jiandong@163.com; Tel: +44784642-4479, +8613895003915; Fax: +8609512066815.

† These two authors contributed equally.

Abstract: *Purpose:* Coronary microvascular dysfunction (CMD) is emerging as an important cause of

myocardial ischemia, but there is a lack of a non-invasive method for reliable early detection of CMD. *Aim:* To develop an electrocardiogram (ECG)-based machine learning algorithm for CMD detection that will lay the groundwork for patient-specific non-invasive early detection of CMD. *Methods:* Vectorcardiography (VCG) was calculated from each 10-second ECG of CMD patients and healthy controls. Sample entropy (*SampEn*), approximate entropy (*ApEn*), and complexity index (*CI*) derived from multiscale entropy were extracted from ST-T segments of each lead in ECGs and VCGs. The most effective entropy subset was determined using the sequential backward selection algorithm under the intra-patient and inter-patient schemes, separately. Then, the corresponding optimal model was selected from eight machine learning models for each entropy feature based on five-fold cross-validations. Finally, the classification performance of *SampEn*-based, *ApEn*-based, and *CI*-based models was comprehensively evaluated and tested on a testing dataset to investigate the best one under each scheme. *Results: ApEn*-based SVM model was validated as the optimal one under the intra-patient scheme, with all testing evaluation metrics over 0.8. Similarly, *ApEn*-based SVM model was selected as the best one under the intra-patient scheme, with major evaluation metrics over 0.8. *Conclusions:* Entropies derived from ECGs and VCGs can effectively detect CMD under both intra-patient and inter-patient schemes. Our proposed models may provide the possibility of an ECG-based tool for non-invasive detection of CMD.

Keywords: coronary microvascular dysfunction (CMD); myocardial ischemia; entropy; machine learning; electrocardiogram (ECG); vectorcardiogram (VCG)

1. Introduction

Clinically, myocardial ischemia refers to a condition in which the perfusion of heart muscle is insufficient, resulting in the reduction of oxygen supply to the heart, abnormal myocardial energy metabolism, and abnormal work of the heart [1]. It can increase cardiovascular events including sudden cardiac death and acute myocardial infarction [2−4], contributes 16% to the world's total deaths, and has been considered as the leading cause of mortality. Obstructive coronary artery disease (CAD), defined as severe stenosis in any epicardial coronary arteries (diameter ~ 5 mm) determined via coronary angiography (CAG, diameter severity [DS] $\geqslant$ 50%) or a fractional flow reserve (FFR < 0.8) [5], had been quantitatively investigated and well established as a main etiology of myocardial ischemia. However, up to 70% of patients undergoing CAG do not have obstructive CAD [6]. Patients with symptoms and signs of ischemic heart disease (IHD) but found to have nonobstructive coronary arteries (INOCA) are increasingly recognized [7]. Coronary microvascular dysfunction (CMD) has been recognized as a major cause of INOCA by the latest European Society of Cardiology guidelines [8,9]. Since the coronary microvasculature (diameter < 500 μm) has remained elusive to conventional medical images [10], the clinical significance of CMD has not been fully understood [11].

CMD refers to impaired blood flow in the coronary microcirculation [11,12]. CMD can be caused by structural or functional abnormalities in the coronary microcirculation, resulting in an inefficacy increase of myocardial blood supply in response to an increased myocardial oxygen demand or coronary microvascular spasm [8,13] and eventually leading to myocardial ischemia. Patients with CMD have higher risks of poor prognosis, re-hospitalization, adverse cardiovascular events, and mortality [9,14]. Recently, the functional assessment or detection of CMD has gained much

popularity due to the substantial proportion of patients developing symptoms and signs of myocardial ischemia, in spite of the absence of obstructive CAD, or after the well-established treatment of obstructive CAD [5]. The prevalence of CMD has reported to be 50 to 65% of angina patients with INOCA [14], and more than 50% in patients with obstructive CAD [15]. CMD is common but easy to be ignored [16]. Therefore, making an early and accurate diagnosis of CMD plays an increasingly critical role in clinical practice. However, the natural course of CMD tends to be relatively long and asymptomatic, making early diagnosis difficult [17], and there is no reliable noninvasive test available.

Currently, the gold standard of diagnostic criteria for CMD is the invasive measurement of index of microcirculatory resistance (IMR) and FFR [6,18−20]. They are not applicable for the early detection of CMD, given that only a minority of inpatient with acute chest pain meets the clinical threshold that requires invasive testing. Besides the invasive measurement of IMR, methods have been developed for the non-invasive evaluation of cardiac microcirculation, including Doppler echocardiography, stress cardiac magnetic resonance imaging, and computational fluid simulation [13]. However, these techniques are not applied as an accessible tool for early detection of CMD in daily clinical practice due to their radiation, high cost, invasiveness, and complicated operation [21]. Since the 2010s, coronary microcirculatory dysfunction has been focused on as a main etiology of myocardial ischemia [20]. There is an increasing clinical demand for a non-invasive, accessible, and cost-efficient tool to recognize patients with a high likelihood of CMD who may subsequently be referred for invasive detection and further diagnosis [22].

In comparison, electrocardiogram (ECG)-based analysis is easy-to-use, low-cost, without the requirement of expertise or expensive devices. CMD can lead to ischemic ECG changes [22,23], which provides the possibility of ECG-based non-invasive detection of CMD. Specially, standard deviation of normal R-R intervals (SDNN) < 100 ms [24], ischemic ST-segment changes [25,26], the prolongation of the heart rate-corrected QT interval [27,28], and variations in T wave [22,27] have been validated as indicators of CMD in patients with INOCA. Sara et al. extracted three features from T wave and developed a linear discriminant classifier which yielded the sensitivity and specificity of 65.9 and 67.7% in detecting CMD [22]. These studies preliminarily validated the effectiveness of ECG-based features in detecting CMD, which provides the possibility of computer-aided CMD detection method using machine learning algorithm. Nevertheless, there are very few computer-aided diagnostic algorithms based on ECG-based features for CMD detection, given that previous studies mainly focused on detecting obstructive CAD [29].

Entropy is a common measure of complexity of time series [30]. Myocardial ischemia-induced variations in waveform and frequency components at certain ECG segments can be reflected in the changes of its entropy. Sample entropy (*SampEn*), approximate entropy (*ApEn*), and multiscale entropy (*MSE*) have been extracted from heart rate variability [31], ST-segments [32], or filtered 12-lead ECGs [33] for evaluating the changes of complexity in ECGs from patients with obstructive CAD [34]. However, the performances of these entropies in detecting CMD are yet to be explored. Therefore, we aim to investigate whether *SampEn*, *ApEn*, and *MSE* extracted from beat-to-beat ST-T segments of each lead in ECG and vectorcardiography (VCG) could reliably detect CMD, and develop an ECG-based CMD detection algorithm which will lay the groundwork for patient-specific, early, and non-invasive detection of CMD by comparing different entropy subsets and machine learning models.

2. Materials and methods

Our work contains five parts: data collection, signal preprocessing, feature extraction, feature selection, and model evaluation and testing, as shown in Figure 1. Firstly, 12-lead ECGs from CMD patients and healthy controls were collected and these from healthy controls were split into 10-second (10 s) segments. After 10 s ECGs were cleansed and denoised, VCGs were synthesized from 12-lead ECGs mathematically. The ST-wave onset and T-wave offset were located on VCGs and ECGs employing a squeeze and regional approach respectively. Three entropy features, i.e., *SampEn*, *ApEn*, and *MSE*-derived complexity index (*CI*), were extracted from the time series obtained by splicing ST-T segments of each lead in ECGs and VCGs beat by beat. Subsequently, the most effective subset of each entropy feature for each model was determined using the sequential backward selection (SBS) algorithm under the intra-patient and inter-patient schemes. Under each scheme, eight different machine learning models for each entropy feature were comprehensively evaluated on a five-fold cross-validation method on the training-validation datasets at the very first place, and then tested on the corresponding testing datasets for every scenario. Finally, the best model for CMD detection was investigated from the three entropy-based models (i.e., *SampEn*-based, *ApEn*-based, and *CI*-based) under each scheme.

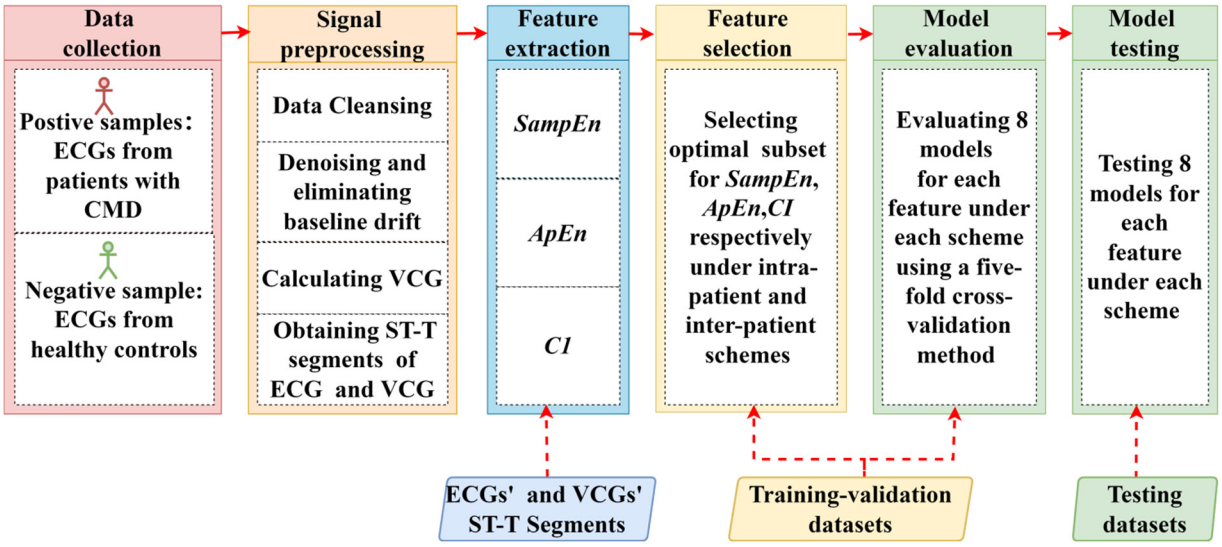

Figure 1. The schematic diagram of our proposed algorithms.

2.1. Data collection

In our study, ECGs and the clinical characteristics of 177 subjects in two cohorts were collected. Regard to positive samples, 99 10 s, 12-lead resting ECG recordings of 70 patients with CMD were obtained from the Second Affiliated Hospital of Zhejiang University, Zhejiang, China, from June 2019 to September 2022 with approval from the local ethics committee for sharing and analyzing retrospective anonymised patient data with informed consent form waived.

Inclusion criteria were described as follows: 1) existence of stable or chronic symptoms of myocardial ischemia; 2) presence of coronary stenosis assessed by CAG but a FFR > 0.8; 3) presence of symptoms of chest pain or tightness but no obvious evidence of myocardial ischemia; 4) presence

of objective evidence of myocardial ischemia.

Exclusion criteria applied to the study population were given by: 1) existence of unsuccessful measurement of FFR or IMR; 2) presence of myocardial infarction in the target vessel within 72 hours; 3) existence of isotonic creatinine > 150 μmol/L, glomerular filtration rate < 45 mL/kg/m^2, or left ventricular ejection fraction < 30%; 4) existence of the following diseases: bundle branch blocks, atrial fibrillation, left ventricular hypertrophy, congestive heart failure, pulmonary arterial hypertension, or heart valve disease.

CMD was diagnosed according to the standardized diagnostic criteria proposed by COVADIS (Coronary Vasomotor Disorders International Study Group): IMR ≥ 25 in the absence of overt obstructive CAD (FFR > 0.80) [12,35], as illustrated in Figure 3. FFR and IMR were measured in the usual fashion by employing a coronary guidewire (St. Jude Medical Inc., PressureWireCertus, C12008) placed in target artery during CAG [35].

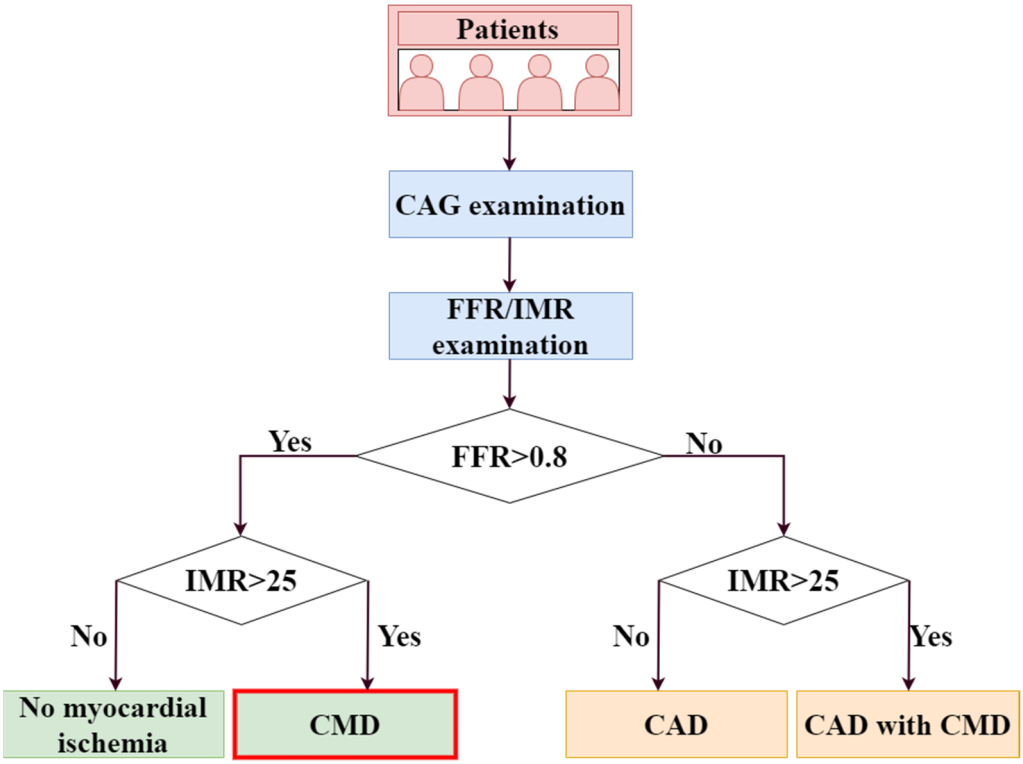

Figure 2. The flowchart of diagnosing CMD.

Finally, patients with CMD were selected whilst patients with CAD were ruled out, as shown in Figure 3. Totally, 99 10 s, 12-lead ECGs (sampling rate: 500 Hz; resolution: 16-bit with 4 μV/LSB) were recorded utilizing a commercially available electrocardiograph (Netherlands Philips Electronics Co. LTD, PageWriter TC30). The clinical characteristics of CMD patients are listed in Table 1.

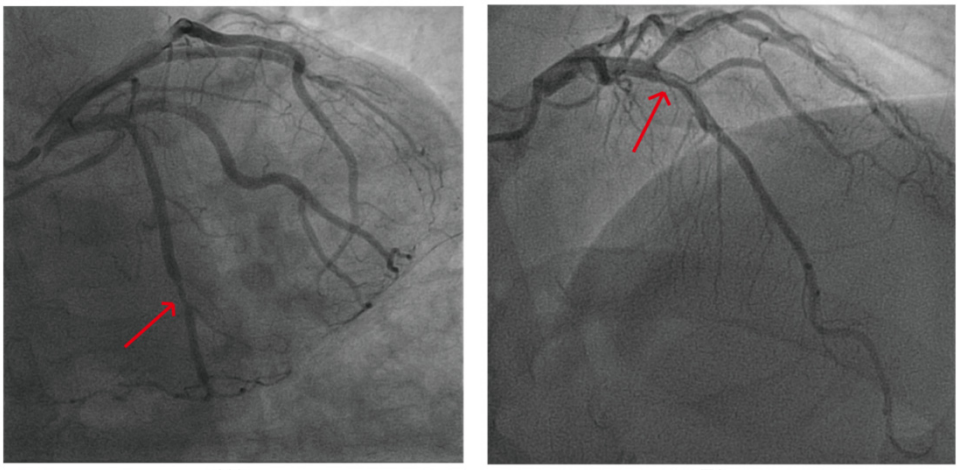

Figure 3. Comparison of CAG between two patients with CAD and CMD. (A) A patient with CAD. There is a coronary stenosis 60% in distal left circumflex artery with FFR < 0.8 and IMR < 25. (B) A patient with CMD. There is a coronary stenosis 40% in left anterior descending artery with IMR > 25 and FFR > 0.8.

Table 1. Clinical characteristics of patients with CMD.

Characteristics	**Value***
Age, years	67.29 ± 9.23
Female, n (%)	26/70
Hypertension, n (%)	32/70
Diabetes mellitus, n (%)	22/70
Heart rate, bpm	70.24 ± 12.87
Systolic blood pressure, mmHg	130.82 ± 15.34
Diastolic blood Pressure, mmHg	73.43 ± 10.45
Smoking history, n (%)	12/70
Family history of CAD, n (%)	12/70

*The categorical data are provided as numbers and percentages, and all of numerical variables are afforded in the form mean ± standard deviation for normally distributed data.

Regarding negative samples, 107 12-lead ECGs (duration: ≥10 seconds; sampling rate: 500 Hz; resolution: 16-bit with 1 mV/LSB) were collected from 107 healthy controls (age: 36.3 ± 15.1 years, 73 females and 34 males) of China Physiological Signal Challenge in 2018 database [36].

2.2. Signal preprocessing

After ECGs from healthy controls were segmented into 10 s ones, data cleansing was implemented to guarantee the quality of ECG signals. Low-quality ECGs resulted from the shedding of the leads or/and poor attachment between the lead and body surface were eliminated, as illustrated in Figure 4. After data cleansing, in total, 301 10 s, 12-lead ECGs (99 CMD, 202 controls) were finally

selected for further analysis.

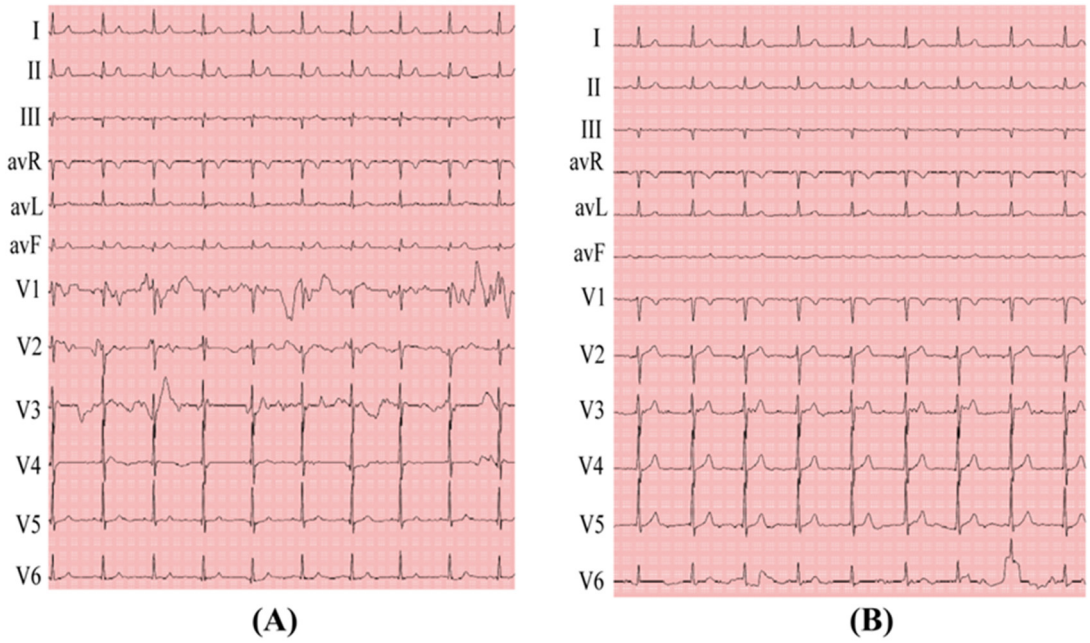

Figure 4. Examples of low-quality ECGs. (A) Leads V1 and V2 have poor attachment to the body surface. (B) The lead avF is shed.

Then, the baseline drift and low-frequency fluctuations were removed using two moving median filters connected in cascade (with length 1.2 and 0.6 seconds respectively) to maximally avoid the deformation of ST segment in comparison with the classical high-pass filter. The output of first median filter was fed into the second one. High-frequency power-line interference and electromyogram noise were eliminated employing discrete wavelet transform (DWT). In detail, Coif4 determined as a wavelet basis function was adopted to decompose noise-containing ECGs up to four levels [37]. After the elimination of noise by an adaptive threshold, the processed signals were transformed back by utilizing the inverse of DWT [37].

Subsequently, all 10 s, 12-lead ECGs were standardized utilizing the clinical standard gain [37,38] and then transformed into VCGs [39], as shown in Figure 5(C−F).

$$\begin{bmatrix} V_x \\ V_y \\ V_z \end{bmatrix} = D \cdot F^T \tag{1}$$

where $D = \begin{bmatrix} 0.38 & -0.07 & -0.13 & 0.05 & -0.01 & 0.14 & 0.06 & 0.54 \\ -0.07 & 0.93 & 0.06 & -0.02 & -0.05 & 0.06 & -0.17 & 0.13 \\ 0.11 & -0.23 & -0.43 & -0.06 & -0.14 & -0.20 & -0.11 & 0.31 \end{bmatrix}$,

$F = [I \quad II \quad V1 \quad V2 \quad V3 \quad V4 \quad V5 \quad V6]$; I, II, $V1$, $V2$, $V3$, $V4$, $V5$, and $V6$ are leads of ECG; V_x, V_y, and V_z are leads of VCG.

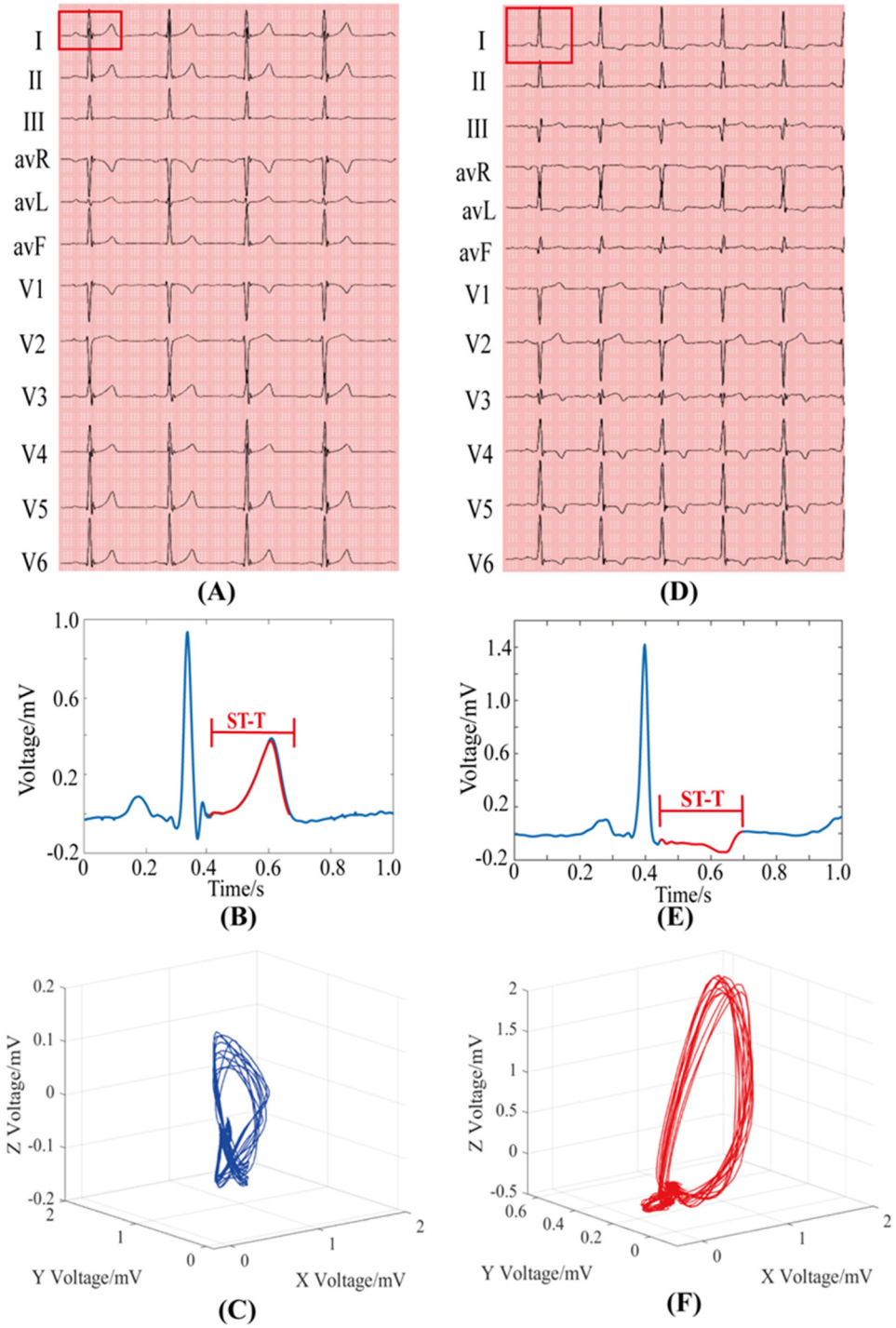

Figure 5. ECG, ST-T segment, and VCG. (A) ECG for a healthy control. (B) ST-T segment for a healthy control. (C) VCG for a healthy control. (D) ECG for a patient with CMD. (E) ST-T segment for a patient with CMD. (F) VCG for a patient with CMD.

Ultimately, the ST-segment onset and T-wave offset were simultaneously marked on the VCGs and ECGs employing a hybrid approach [40] to obtain ST-T segments for calculating entropies, as shown in Figure 5(B−E). ST-wave onsets were located using a squeeze algorithm, and T-wave offsets using a regional approach [40].

2.3. Entropy feature extraction

Sample entropy (*SampEn*), approximate entropy (*ApEn*), and multiscale entropy (*MSE*) have been demonstrated to be electrocardiographic indictors for evaluating the changes of complexity in ECGs from patients with obstructive CAD [33,34]. Similarly, *SampEn*, *ApEn*, and *MSE*-derived *CI* were calculated from ST-T segments of each VCG and ECG lead to assess the changes of complexity in ECGs induced by CMD.

For each lead in VCGs or ECGs, its time series, i.e., $\{t(l)\} = t(1), t(2), ..., t(L)$ was obtained by splicing ST-T segments beat by beat and standardized as follows:

$$t(l) = \frac{t(l) - \mu}{\sigma} \tag{2}$$

where L presents the length of the time series; σ and μ are the standard deviation and mean value of the time series.

2.3.1. SampEn

SampEn is defined as the negative natural logarithm of a conditional probability [30] and calculated to measure the complexity of beat-to-beat ST-T segment in this work. For each VCG and ECG lead, *SampEn* ($S_i, i = V_x, V_y, V_z, I, II, III, avR, avL, avF, V1, V2, V3, V4, V5, V6$) was calculated from the time series, i.e., $\{t(l)\} = t(1), t(2), ..., t(L)$ as follows [37,41]:

$$SampEn(L, m, \varepsilon) = -\ln \frac{\psi^{m+1}(\varepsilon)}{\psi^{m}(\varepsilon)} \tag{3}$$

where $\psi^{m}(\varepsilon)$ and $\psi^{m+1}(\varepsilon)$ are the average probability that two vectors coincide for m and $m+1$ points, respectively; ε denotes the tolerance for accepting matches, and m refers to embedded dimension [30].

In this work, the value of ε was set to 0.1, and that of m was set to 2 [30,37].

2.3.2. ApEn

In our work, *ApEn* ($A_i, i = V_x, V_y, V_z, I, II, III, avR, avL, avF, V1, V2, V3, V4, V5, V6$) of time series $\{t(l)\} = t(1), t(2), ..., t(L)$ of each lead in ECGs and VCGs were calculated [30,41].

Firstly, the vectors $T_m(1), ..., T_m(L-m+1)$ and $T_{m+1}(1), ..., T_{m+1}(L-m)$ with a dimension of m and $m+1$ were formed to be $T_m(z) = \{t(z), t(z+1), ..., t(z+m-1)\}, 1 \le z \le L-m+1$ and $T_{m+1}(z) = \{t(z), t(z+1), ..., t(z+m)\}, 1 \le z \le L-m$.

Then, the distances between any $T_m(z)$ and $T_m(p)$, and between any $T_{m+1}(z)$ and $T_{m+1}(p)$ were defined as:

$$D\left[T_m(z),T_m(p)\right]=\max_{0\le q\le m-1}\left(\left|t(z+q)-t(p+q)\right|\right),1\le z,p\le L-m+1 \tag{4}$$

$$D\left[T_{m+1}(z),T_{m+1}(p)\right]=\max_{0\le q\le m}\left(\left|t(z+q)-t(p+q)\right|\right),1\le z,p\le L-m \tag{5}$$

Subsequently, $C_j^m(\varepsilon)$ and $C_j^{m+1}(\varepsilon)$ were given by:

$$C_z^m(\varepsilon)=\frac{u_z^m(\varepsilon)}{L-m+1},z=1,\cdots,L-m+1 \tag{6}$$

$$C_z^{m+1}(\varepsilon)=\frac{u_z^{m+1}(\varepsilon)}{L-m},z=1,\cdots,L-m \tag{7}$$

where $u_z^m(\varepsilon)$ denotes the number that the distance between any vector $T_m(z)$ and vector $T_m(p)$ is within ε, and similarly $u_z^{m+1}(\varepsilon)$ is the number that the distance between any vector $T_{m+1}(z)$ and vector $T_{m+1}(p)$ is within ε. ε denotes the tolerance for accepting matches [30].

$$\theta^m(\varepsilon)=\frac{\sum_{z=1}^{L-m+1}\log C_z^m(\varepsilon)}{L-m+1} \tag{8}$$

$$\theta^{m+1}(\varepsilon)=\frac{\sum_{z=1}^{L-m}\log C_z^{m+1}(\varepsilon)}{L-m} \tag{9}$$

where $\theta^m(\varepsilon)$ is the average probability that two any vectors coincide for $m+1$ points, and $\theta^{m+1}(\varepsilon)$ is the average probability that two any vectors coincide for $m+1$ points.

Ultimately, *ApEn* was calculated as follows:

$$ApEn(\varepsilon,m)=\theta^m(\varepsilon)-\theta^{m+1}(\varepsilon) \tag{10}$$

In this work, the value of ε was set to 0.1 and that of m was set to 2 [30,37].

2.3.3. CI

MSE builds on *SampEn* algorithm by integrating a coarse-graining procedure and provides insight into the point-to-point fluctuation across multiple time scales [42]. The kth coarse-grained time series for each time scale of τ, $y^{(\tau)}=\left\{y_1^{(\tau)},y_2^{(\tau)},\cdots,y_p^{(\tau)}\right\}$, was defined as [42]:

$$y_k^{(\tau)}=\frac{1}{\tau}\sum_{l=(k-1)\tau+1}^{k\tau-1}t(l),1\le k\le\frac{L}{\tau} \tag{11}$$

where $p = \frac{L}{\tau}$, τ is the time scale of interest, L is the length of the time series, and $t(k)$ is a data point in the time series, i.e., $\{t(l)\} = t(1), t(2), ..., t(L)$, of each ECG and VCG lead.

SampEn algorithm was employed to obtain an entropy value at each time scale.

$$MSE(\tau, m, \varepsilon) = SampEn\left(y^{(\tau)}, m, \varepsilon\right), 1 \le \tau \le M \tag{12}$$

where M is the maximum time scale; the value of ε was set to 0.1 and that of *m* set to 2 [30,37].

A major concern when employing *MSE* algorithm is to guarantee that the data length of the coarse-grained time series is long enough at the maximum time scale for the reason that too short time series at any time scale associates with inconsistent probabilities from *SampEn*. More than 200 data points per time scale has been recommended to elicit consistent *SampEn* values [43]. Since the data length of time series is between 900 and 1200, the maximum time scale was set to 3. The values of τ were set to 1, 2, 3 [30,43].

Once all parameters and *MSE* have been determined, the area under the *MSE* vs. time scale curve, entitled as *CI* [42] of ECGs' and VCGs' ST-T segments were calculated as follows:

$$CI = \sum_{\tau=1}^{M} MSE(\tau) \tag{13}$$

where $MSE(\tau)$ is the *MSE* value at time scale factor of τ.

2.4. CMD detection using machine learning models

2.4.1. The proposed models

Eight common machine learning models [29] i.e., support vector machine (SVM), multilayer perceptron (MLP), gradient boosting [44], bagging ensemble (Bagging), random forest, ada boost, decision tree, and K-nearest neighbors (KNN) were conducted to distinguish between healthy controls and patients with CMD in this work. All machine learning models were implemented in Python 3.7 employing open-source libraries: Scikit-learn (0.20.1, http://scikit-learn.org/stable).

The entropy subsets selected by utilizing the SBS algorithm were fed into each framework to identify the subjects (i.e., CMD patients vs. healthy controls). Regard to the model design, *SampEn*-based model denoted each framework fed with VCGs' and ECGs' *SampEn* only. Similarly, *ApEn*-based and *CI*-based models presented respectively the frameworks utilizing *ApEn* and *CI* only.

2.4.2. Entropy feature selection

Feature selection is widely implemented to determine the optimal feature subset from a large initial set of features thus improving generalization of the machine learning model and reducing the computational demand. To achieve the most effective entropy subset, SBS algorithm [45,46] has been conducted, as presented in Figure 6.

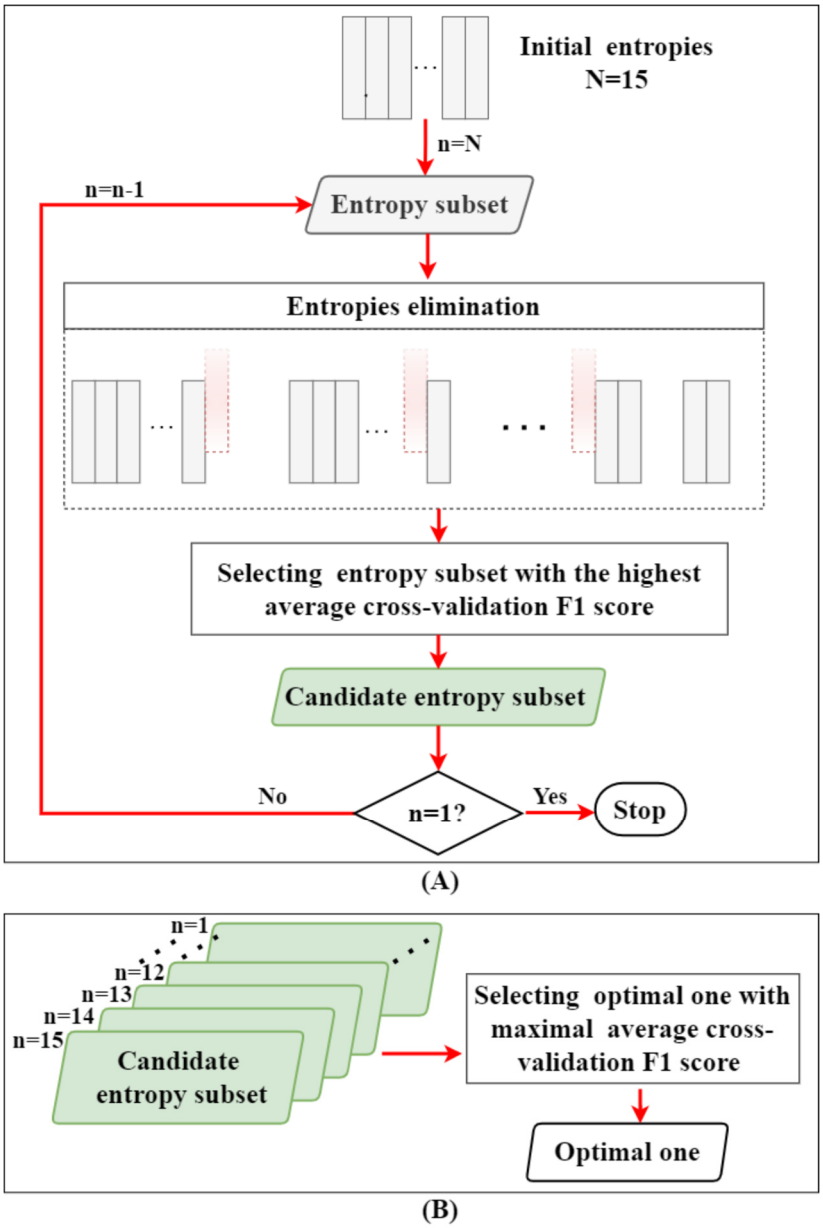

Figure 6. The flowchart of the SBS algorithm for determining the optimal entropy feature subset. (A) The entropy elimination technique. (B) The selection of the optimal entropy subset. The SBS algorithm was conducted on *SampEn*, *ApEn*, and *CI* separately; where the initial entropies consist of all the features from 12 ECG and 3 VCG leads. In each iteration step, the candidate subset (each with n-1 elements) was selected from n subsets of the parent set with n elements. The selected candidate subset was the parent set in the next iteration.

First, a collection of candidate entropy subset for each model was created on the training-validation dataset, as illustrated in Figure 6A. For each algorithm, a baseline model was constructed using each entropy feature separately, i.e., 15 *SampEn*, 15 *ApEn*, or 15 *CI* calculated from ST-T segments of 3-lead VCGs and 12-lead ECGs. The iteration variable n was initially set as 15. Each subset with n-1 entropies was trained and evaluated by employing a five-fold cross-validation to select the one with the highest average F1 score, which takes into account precision and recall, and is not sensitive to the class-imbalance [47] into next iteration where subsets with n-2 entropies derived from

the selected n-1 element subset were evaluated and compared similarly. The iteration continued until n = 1. Finally, 15 candidate subsets with different numbers of entropies (i.e., 1, 2,…, 15) were selected by the iterative SBS algorithm.

Second, the optimal entropy subset, which had maximal five-fold cross-validation average F1 score, was identified from the SBS-generated collection of the candidate entropy subsets achieved above, as illustrated in Figure 6B.

Finally, the optimal entropy subset was separately determined from the total 15 *SampEn* for each *SampEn*-based model. Similarly, optimal entropy subsets for *ApEn*-based and *CI*-based models were determined.

2.5. Experimental methodology for model evaluation and testing

The utility of the constructed models was evaluated by accuracy, specificity, sensitivity, and F1 score [37]. Additionally, the geometric mean (G-mean) and an area under the curve for precision/recall (PR-AUC) were considered as evaluation metrics. The geometric mean (G-mean) is the geometric mean of sensitivity and specificity [48]. PR-AUC is a metric commonly used in highly imbalanced datasets [49].

Both intra-patient and inter-patient schemes were implemented to verify the effectiveness of our proposed modes. Under the intra-patient scheme, positive and negative ECGs were randomly split in 80% for training and validation, and 20% for testing at the same time. Regarding the inter-patient scheme, patients were randomly separated in the proportion of 4:1 for training-validation and testing, and the corresponding ECGs formed the training-validation and testing datasets. Under both scheme, a five-fold cross-validation approach was adopted in the training-validation dataset to optimize model parameters and check evaluation metrics (i.e., sensitivity, accuracy, specificity, F1 score, G-mean, and PR-AUC) on the training-validation dataset for each model at the very first place. Then, each trained model was tested on the testing dataset to check whether the selected model always performed best under each scheme for every scenario.

In this work, the signal preprocessing and feature extraction were developed utilizing MATLAB (R2021; The MathWorks Inc. Natick, USA). All machine learning models were established, evaluated, and tested in Python 3.7 as well as feature selection.

3. Results

3.1. Intra-patient scheme

The evaluation metrics of *SampEn*-based, *ApEn*-based, and *CI*-based models on the validation dataset are illustrated in first, second, and third columns of Figure 7 respectively. It can be observed from the first column that SVM model outperforms other *ApEn*-based models in response to the major evaluation metrics. Therefore, SVM was determined as the candidate one for *SampEn*-based model. Similarly, MLP and SVM models were selected as the candidates for *SampEn*-based and *CI*-based models respectively. As listed in Table 2, the testing evaluation metrics of *ApEn*-based SVM model are higher than those of other *ApEn*-based models. Similarly, *SampEn*-based MLP and *CI*-based SVM models provide the highest the testing evaluation metrics across all *SampEn*-based and *CI*-based ones respectively.

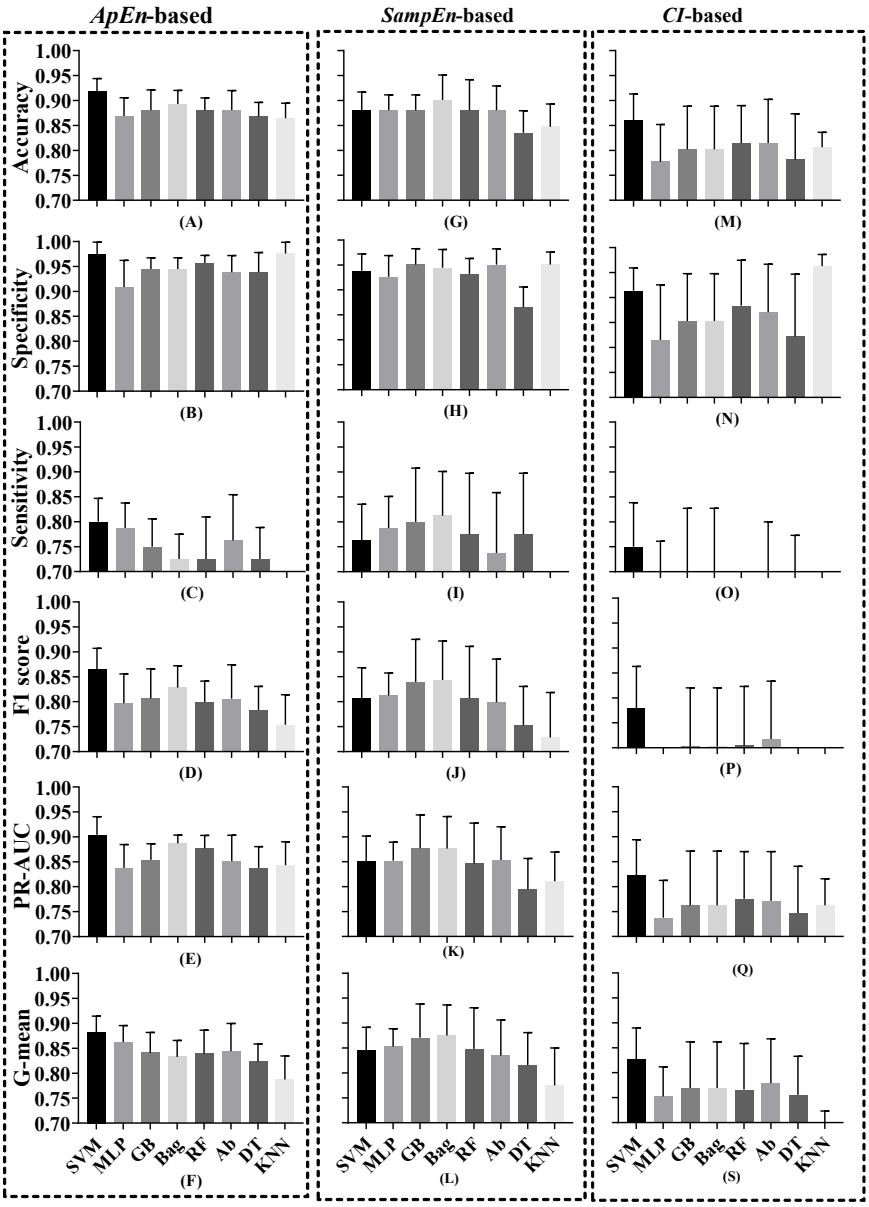

Figure 7. The comparison of the classification performance between different models under the intra-patient scheme. (A) Accuracy of *ApEn*-based models. (B) Specificity of *ApEn*-based models. (C) Sensitivity of *ApEn*-based models. (D) F1 score of *ApEn*-based models. (E) PR-AUC of *ApEn*-based models. (F) G-mean of *ApEn*-based models. (G) Accuracy of *SampEn*-based models. (H) Specificity of *SampEn* -based models. (I) Sensitivity of *SampEn* -based models. (J) F1 score of *SampEn*-based models. (K) PR-AUC of *SampEn*-based models. (L) G-mean of *SampEn* -based models. (M) Accuracy of *CI*-based models. (N) Specificity of *CI*-based models. (O) Sensitivity of *CI*-based models. (P) F1 score of *CI*-based models. (Q) PR-AUC of *CI*-based models. (S) G-mean of *CI*-based models. GB: Gradient Boost; Bag: Bagging; RF: Random Forest; Ab: Ada boost; DT: Decision Tree. G-mean: geometric mean; PR-AUC: an area under the curve for precision/recall.

Subsequently, the comparisons among the three candidate models (i.e., *ApEn*-based SVM, *SampEn*-based MLP, and *CI*-based SVM models) show that *ApEn*-based SVM model affords the

highest evaluation metrics on the validation dataset, as shown in Figure 7. Furthermore, the classification performances of the three candidate models on the testing dataset present that *ApEn*-based SVM model is superior to the other two models (i.e., *SampEn*-based MLP and *CI*-based SVM models) with all evaluation metrics higher than 0.8, as listed in Table 2. Thus, *ApEn*-based SVM model was validated as the optimal one in detecting CMD under the intra-patient scheme.

Table 2. Comparisons of the testing evaluation metrics of *ApEn*-based, *SampEn*-based, and *CI*-based models under the intra-patient scheme.

Models	Accuracy	Specificity	Sensitivity	F1 score	PR-AUC	G-mean
***ApEn*-based SVM**	**0.898**	**0.925**	**0.842**	**0.842**	**0.868**	**0.883**
ApEn-based MLP	0.881	0.925	0.789	0.811	0.845	0.855
ApEn-based Gradient Boost	0.729	0.750	0.684	0.619	0.676	0.716
ApEn-based Bagging	0.881	0.950	0.737	0.800	0.848	0.837
ApEn-based Random Forest	0.847	0.950	0.632	0.727	0.804	0.775
ApEn-based Ada Boost	0.814	0.900	0.632	0.686	0.750	0.754
ApEn-based Decision Tree	0.847	0.900	0.737	0.757	0.800	0.814
ApEn-based KNN	0.831	0.925	0.632	0.706	0.775	0.764
SampEn-based SVM	0.831	0.825	0.842	0.762	0.794	0.833
***SampEn*-based MLP**	**0.881**	**0.925**	**0.789**	**0.811**	**0.845**	**0.855**
SampEn-based Gradient Boost	0.814	0.850	0.737	0.718	0.761	0.791
SampEn-based Bagging	0.847	0.900	0.737	0.757	0.800	0.814
SampEn-based Random Forest	0.814	0.875	0.684	0.703	0.754	0.774
SampEn-based Ada Boost	0.847	0.900	0.737	0.757	0.800	0.814
SampEn-based Decision Tree	0.746	0.750	0.737	0.651	0.702	0.743
SampEn-based KNN	0.712	0.825	0.476	0.514	0.603	0.625
***CI*-based SVM**	**0.847**	**0.875**	**0.789**	**0.769**	**0.804**	**0.831**
CI-based MLP	0.831	0.875	0.737	0.737	0.779	0.803
CI-based Gradient Boost	0.746	0.775	0.684	0.634	0.688	0.728
CI-based Bagging	0.780	0.825	0.684	0.667	0.718	0.751
CI-based Random Forest	0.814	0.875	0.684	0.703	0.754	0.773
CI-based Ada Boost	0.746	0.800	0.632	0.615	0.675	0.711
CI-based Decision Tree	0.746	0.750	0.737	0.651	0.702	0.743
CI-based KNN	0.712	0.825	0.474	0.514	0.603	0.625

Note: Best performance for *ApEn*-based, *SampEn*-based, and *CI*-based models is highlighted in bold.

3.2. Inter-patient scheme

Figure 8 illustrates the comparisons of evaluation metrics among *ApEn*-based, *SampEn*-based, and *CI*-based models on the validation dataset for detecting CMD. From the first column, SVM model yields the highest classification performance across all *ApEn*-based models. Therefore, it was verified as the candidate one for *ApEn*-based model. Similarly, Bagging and SVM models were determined as the candidates for *ApEn*-based and *CI*-based models respectively. Regard to the classification performances on the testing dataset, *ApEn*-based SVM model outperforms other *ApEn*-based ones, as

listed in Table 3. Similarly, *SampEn*-based Bagging and *CI*-based SVM models perform best across all *SampEn*-based and *CI*-based ones respectively.

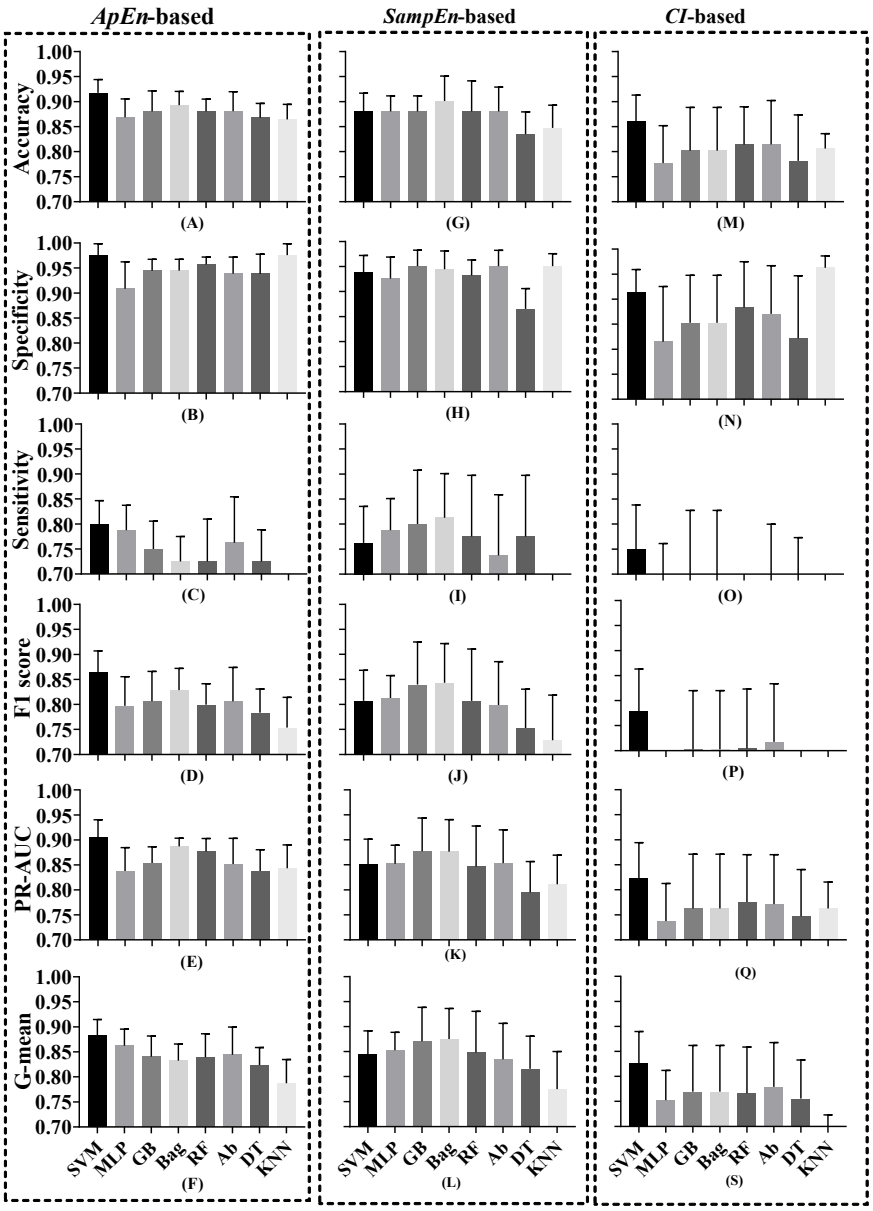

Figure 8. The comparison of classification capability between different models under the inter-patient scheme. (A) Accuracy of *ApEn*-based models. (B) Specificity of *ApEn*-based models. (C) Sensitivity of *ApEn*-based models. (D) F1 score of *ApEn*-based models. (E) PR-AUC of *ApEn*-based models. (F) G-mean of *ApEn*-based models. (G) Accuracy of *SampEn*-based models. (H) Specificity of *SampEn* -based models. (I) Sensitivity of *SampEn* -based models. (J) F1 score of *SampEn*-based models. (K) PR-AUC of *SampEn*-based models. (L) G-mean of *SampEn* -based models. (M) Accuracy of *CI*-based models. (N) Specificity of *CI*-based models. (O) Sensitivity of *CI*-based models. (P) F1 score of *CI*-based models. (Q) PR-AUC of *CI*-based models. (S) G-mean of *CI*-based models. GB: Gradient Boost; Bag: Bagging; RF: Random Forest; Ab: Ada boost; DT: Decision Tree. G-mean: geometric mean; PR-AUC: an area under the curve for precision/recall.

Then, the classification performances of *SampEn*-based Bagging, *ApEn*-based SVM, and *CI*-based SVM models were compared to select the optimal one. It can be observed from Figure 8 that *ApEn*-based SVM model has superiority to the other two models except the sensitivity for *SampEn*-based Bagging model. Moreover, the comparisons of the three candidate models on the testing dataset demonstrate that *ApEn*-based SVM model has all testing evaluation metrics higher than *SampEn*-based Bagging and *CI*-based SVM models, as listed in Table 3. Hence, *ApEn*-based SVM model was verified as the optimal one for CMD detection under the inter-patient scheme.

Table 3. Comparisons of the testing evaluation metrics of *ApEn*-based, *SampEn*-based, and *CI*-based models under the inter-patient scheme.

Models	Accuracy	Specificity	Sensitivity	F1 score	PR-AUC	G-mean
***ApEn*-based SVM**	**0.881**	**0.925**	**0.789**	**0.811**	**0.845**	**0.855**
ApEn-based MLP	0.814	0.775	0.895	0.756	0.791	0.833
ApEn-based Gradient Boost	0.797	0.850	0.684	0.684	0.735	0.763
ApEn-based Bagging	0.864	0.925	0.737	0.778	0.823	0.826
ApEn-based Random Forest	0.847	0.925	0.684	0.743	0.799	0.796
ApEn-based Ada Boost	0.847	0.925	0.684	0.743	0.799	0.796
ApEn-based Decision Tree	0.847	0.925	0.684	0.743	0.799	0.796
ApEn-based KNN	0.831	0.925	0.632	0.706	0.775	0.764
SampEn-based SVM	0.847	0.900	0.737	0.757	0.800	0.814
SampEn-based MLP	0.814	0.850	0.737	0.718	0.761	0.791
SampEn-based Gradient Boost	0.864	0.900	0.789	0.789	0.823	0.843
***SampEn*-based Bagging**	**0.864**	**0.925**	**0.737**	**0.778**	**0.823**	**0.826**
SampEn-based Random Forest	0.864	0.925	0.737	0.778	0.823	0.826
SampEn-based Ada Boost	0.831	0.925	0.632	0.706	0.775	0.764
SampEn-based Decision Tree	0.797	0.850	0.684	0.684	0.735	0.763
SampEn-based KNN	0.814	0.900	0.632	0.686	0.750	0.754
***CI*-based SVM**	**0.831**	**0.875**	**0.737**	**0.737**	**0.779**	**0.803**
CI-based MLP	0.729	0.750	0.684	0.619	0.676	0.716
CI-based Gradient Boost	0.763	0.825	0.632	0.632	0.691	0.722
CI-based Bagging	0.763	0.825	0.632	0.632	0.691	0.722
CI-based Random Forest	0.729	0.825	0.526	0.556	0.633	0.659
CI-based Ada Boost	0.695	0.750	0.579	0.550	0.619	0.659
CI-based Decision Tree	0.678	0.675	0.684	0.578	0.643	0.680
CI-based KNN	0.780	0.950	0.421	0.552	0.704	0.632

Note: Best performance for *ApEn*-based, *SampEn*-based, and *CI*-based models is highlighted in bold.

4. Discussion

4.1. Comparison with existing studies

The *ApEn*-based SVM models were verified to be the optimal ones for CMD detection, with all and major testing evaluation metrics over 0.8 under the intra-patient and inter-patient schemes

respectively, indicating the feasibility of entropy-based early detection of CMD using machine learning algorithm.

Obstructive CAD and CMD are recognized as two major causes of myocardial ischemia. Since obstructive CAD had been considered as the leading cause of myocardial ischemia, previous studies of ECG-based machine learning algorithms mainly focused on obstructive CAD detection [29,50], with a lack of in-depth investigation in patients with CMD. An advantage of the ECG-based method is its interpretability. Some mathematical models disclosed the pathology of CMD from cell levels. Some critical electrophysiological phenomena and parameters, e.g., T wave alternans, can be directly reflected on ECG and VCG, providing the possibility of fine-grained classification and stratification of CMD [51,52]. To detect CMD non-invasively, electrocardiographic indictors., i.e., SDNN < 100 ms [24], ischemic ST-T segment changes [22,25−27], and the prolongation of the heart rate-corrected QT interval [27,28] have been investigated. Specially, SARA et al. combined three features, i.e., T-wave area, T Peak-T end, and Y-center of gravity as inputs of a linear discriminant analysis and recognized the existence presence of an abnormal coronary flow reserve (CFR) with the highest accuracy of 66.5 ± 0.3% and 74 ± 2% for females and males [22]. Since CFR assesses the entire coronary bed including both epicardial and microvascular districts, the authors utilized an abnormal CFR and DS < 50% as the standard for selecting CMD patients to exclude the influence of obstructive CAD. In contrast, IMR used in our work is not influenced by epicardial artery stenosis and can accurately reflect the microvascular flow resistance [18,21].

As far as we known, this is the first machine learning model employing entropy features of ECG and VCG to detect CMD based on reliable clinical diagnosis derived from patient specific FFR and IMR values.

4.2. Relationship between CMD and entropy features

CMD is believed to result from inadequate relaxation of vascular smooth muscle, which mediates myocardial ischemia [1]. During myocardial ischemia, the decrease of conduction velocity and the duration of action potential lead to the heterogeneous repolarization process [53]. Further, ischemia-induced repolarization dispersion causes ischemic changes in ECG and VCG signals. Regarding ECG, the variations in ST segment include ST-segment depression or elevation, and those in T wave include high-tip T wave, biphasic T wave or inverted T wave [51,52], as illustrated in Figure 1. As for VCG, the changes exist in the QRS loop, the spatial orientations and magnitude of ST vectors [51,54], the T-loop morphology, and the T-vector angle [55]. The beat-to-beat variations in T-loop roundness present intrinsic measures of beat-to-beat repolarization ability [55]. Therefore, the variations in ST-T segments of VCGs and ECGs can be considered as indicators of myocardial ischemia [37,51,55].

The ischemic variations have been demonstrated to afford detectable morphology variability in beat-to-beat ST-T segments, while the morphology shows a consistent pattern in normal subjects [56]. CMD can lead to changes in ST segment or/and T wave [22]. The complexity resulting from the morphology variations can be measured by entropy.

For the first time, we observed that CMD-induced complexity changes of beat-to-beat ST-T segments can be reliably detected by *SampEn*, *ApEn*, and *MSE*-derived *CI* (Tables 2 and 3), which provided reference for other ECG-based CMD detection algorithms. Specially, *ApEn* outperforms *SampEn* and *CI* in detecting CMD under both intra-patient and inter-patient schemes. *ApEn* values extracted from pathologic time series are different from those obtained from free running physiologic

systems under healthy condition [57]. In this work, *ApEn* could be used to effectively measure the changed of the complexity induced by CMD.

4.3. The comparison of different models

An SVM model has been widely applied in the detection of myocardial ischemia location and multiclass myocardial infraction classification [50]. The other seven models also have been applied for CAD detection and yield excellent classification performance [29].

We evaluated and tested these common machine learning models to determine the optimal one for CMD detection. Our results present that an SVM model yields the best classification performance across all models under both schemes. Moreover, compared with MLP, gradient boosting, Bagging, and Ada boost models, an SVM model does not require extensive computational resources to determine the classification results. Therefore, it is more accomplishable on wearable devices, given that their computational resources are restricted.

We observed a difference between entropy features in the performance of SVM algorithm in detecting CMD. Under the intra-patient scheme, *ApEn*-based SVM model employing (A_I, A_{II}, A_{avR}, A_{avL}, A_{avF}, A_{V4}, A_{V5}, A_{V6}, A_{Vx}, A_{Vz}) was demonstrated to surpass the other models. Under the inter-patient scheme, *ApEn*-based SVM model utilizing (A_I, A_{II}, A_{III}, A_{V4}, A_{V5}, A_{V6}, A_{Vx}, A_{Vy}) provided the best performance across all models. The patient-specific algorithm optimization deserves further investigation.

4.4. Strengths, limitations, and future directions

The results in this work have demonstrated that *SampEn*, *ApEn*, and *CI* calculated from ST-T segments of ECGs and VCGs could be useful features to recognize a patient with CMD for the first time. Furthermore, our proposed models afford the possibility of noninvasive and cost-effective detection of CMD. It could be easily deployed on conventional ECG acquisition equipment, given that VCG can be calculated from 12-lead ECG by employing a Kors transformation matrix [39]. Additionally, it is based on the 10 s ECG which has been widely accepted as the first-line diagnostic technique in assessing patients with suspected myocardial ischemia due to its non-invasiveness, non-radiation, and low cost [58]. The data length of the time series consisting of beat-to-beat ST-T segments in each lead is between 900 and 1200 and longer than 200 which is suggested for calculating *ApEn* and *SampEn* [43]. The minimum data length is more than 300 with the maximal time scale factor of $\tau = 3$ regarding *MSE*. The commonly used 10 s ECG is available for calculating reliably the three entropy features. Our algorithm meets the clinical needs with no extra workload for operators, given that our algorithms are fully automatic. Hence, our results open a new pathway towards non-invasive, low-cost, and operator-friendly detection of CMD in various application scenarios using wearable ECG sensors. Therefore, the proposed method may provide an approach for large-scale screening of CMD in populations, especially in low-resource areas.

However, some limitations exist in our work. First, the sample size is small, especially the positive cases. The sample size is limited by the number of patients diagnosed with CMD who underwent full examination. The scarce of CMD patients' data is a common limitation of existing studies (e.g., 41 in [24]; 20, in [59]). Especially, the in vivo measurement of IMR is performed invasively by using an intravascular guide wire with the injection of saline into coronary artery multiple times and is

expensive. Therefore, it is not widely used in clinical practice and only performed for the patients with obvious symptoms of CMD, which limited the data size to develop highly reliable algorithms that suit different cohorts. Secondly, the positive (i.e., CMD patients) and negative (i.e., healthy controls) samples were collected from different cohorts. Hence, the difference in physiological characteristics may influence the classification performance. Finally, only ECG-based features were selected as inputs, whereas clinical characteristic and other medical images were not integrated in our algorithms.

In future studies, multicenter large-scale studies can extend the positive sample size and further verify our results, optimize the algorithm, and improve the classification performance by considering the difference in physiological features among different cohorts. Based on large datasets, the combination between ECG-based features and clinic-characteristics could be considered to boost the improvement in the classification performance. The fine-grained classification of CMD based on ECG-VCG features also deserves further exploration.

5. Conclusions

In conclusion, *SampEn*, *ApEn*, and *CI* extracted from the beat-to-beat ST-T segment of ECGs and VCGs could be useful entropy features to recognize electrocardiographic signatures of CMD under both intra-patient and inter-patient schemes. Our proposed models may provide the possibility of an ECG-based tool for non-invasive detection of CMD.

Acknowledgments

The authors would like to appreciate all participants for data collection and providing constructive comments. This word was funded by the Natural Science Foundation of Ningxia Province (No. 2022AAC03242), the North Minzu University Scientific Research Projects (No. 2021JCYJ10), the Natural Science Foundation of China (NSFC) (No. 62171408), the Major Scientific Project of Zhejiang Lab (No. 2020ND8AD01), the Key Research and Development Program of Zhejiang Province (No. 2020C03016), Ningxia First-Class Discipline and Scientific Research Projects (Electronic Science and Technology) (No. NXYLXK2017A07), Innovation Team of Lidar Atmosphere Remote Sensing of Ningxia Province, the high level talent selection and training plan of North Minzu University, and Plan for Leading Talents of the State Ethnic Affairs Commission of the People's Republic of China.

Conflict of interest

The authors declare there is no conflict of interest.

References

1. P. Severino, A. D'Amato, M. Pucci, F. Infusino, F. Adamo, L. I. Birtolo, et al., Ischemic heart disease pathophysiology paradigms overview: from paque activation to microvascular dysfunction, *Int. J. Mol. Sci.*, **21** (2020), 8118. https://doi.org/10.3390/ijms21218118
2. A. E. Moran, M. H. Forouzanfar, G. A. Roth, G. A. Mensah, M. Ezzati, C. J. L. Murray, et al., Temporal trends in ischemic heart disease mortality in 21 world regions, 1980 to 2010, *Circulation*, **129** (2014), 1483−1492. https://doi.org/10.1161/CIRCULATIONAHA.113.004042

3. L. Dai, Y. Zang, G. Shou, L. Xia, Simulation of MCG signal in 2D cardiac tissue sheet with ischemic condition, in *2011 Computing in Cardiology*, (2011), 21−24. Available from: http://cinc.mit.edu/archives/2011/pdf/0021.pdf.
4. J. Zhang, H. Li, Z. Pu, H. Liu, T. Huang, H. Cheng, et al., Early diagnosis of coronary microvascular dysfunction by myocardial contrast stress echocardiography, *Math. Biosci. Eng.*, **20** (2023), 7845−7858. https://doi.org/10.3934/mbe.2023339
5. F. Mangiacapra, M. M. Viscusi, G. Verolino, L. Paolucci, A. Nusca, R. Melfi, et al., Invasive assessment of coronary microvascular function, *J. Clin. Med.*, **11** (2021), 228. https://doi.org/10.3390/jcm11010228
6. V. Kunadian, A. Chieffo, P. G. Camici, C. Berry, J. Escaned, A. H. E. M. Maas, et al., An EAPCI expert consensus document on Ischaemia with non-obstructive coronary arteries in collaboration with European society of cardiology working group on coronary pathophysiology & microcirculation endorsed by coronary vasomotor disorders international study group, *Eur. Heart J.*, **41** (2020), 3504−3520. https://doi.org/10.1093/eurheartj/ehaa503
7. S. H. Lee, D. Shin, J. M. Lee, T. P. van de Hoef, D. Hong, K. H. Choi, et al., Clinical relevance of Ischemia with nonobstructive coronary arteries according to coronary microvascular dysfunction, *J. Am. Heart Assoc.*, **11** (2022), e025171. https://doi.org/doi:10.1161/JAHA.121.025171
8. J. C. Kaski, F. Crea, B. J. Gersh, P. G. Camici, Reappraisal of ischemic heart disease, *Circulation*, **138** (2018), 1463−1480. https://doi.org/10.1161/CIRCULATIONAHA.118.031373
9. M. G. del Buono, R. A. Montone, M. Camilli, S. Carbone, J. Narula, C. J. Lavie, et al., Coronary microvascular dysfunction across the spectrum of cardiovascular diseases: JACC state-of-the-art review, *J. Am. Coll. Cardiol.*, **78** (2021), 1352−1371. https://doi.org/10.1016/j.jacc.2021.07.042
10. J. Zhou, Y. Onuma, S. Garg, N. Kotoku, S. Kageyama, S. Masuda, et al., Angiography derived assessment of the coronary microcirculation: is it ready for prime time? *Expert Rev. Cardiovasc. Ther.*, **20** (2022), 549−566. https://doi.org/10.1080/14779072.2022.2098117
11. C. N. B. Merz, C. J. Pepine, M. N. Walsh, J. L. F. P. G. Camici, W. M. Chilian, J. A. Clayton, et al., Ischemia and no obstructive coronary artery disease (INOCA), *Circulation*, **135** (2017), 1075−1092. https://doi.org/10.1161/CIRCULATIONAHA.116.024534
12. P. Ong, P. G. Camici, J. F. Beltrame, F. Crea, H. Shimokawa, U. Sechtem, et al., International standardization of diagnostic criteria for microvascular angina, *Int. J. Cardiol.*, **250** (2018), 16−20. https://doi.org/10.1016/j.ijcard.2017.08.068
13. B. Tjoe, L. Barsky, J. Wei, B. Samuels, B. Azarbal, C. N. B. Merz, et al., Coronary microvascular dysfunction: considerations for diagnosis and treatment, *Cleve. Clin. J. Med.*, **88** (2021), 561−571. https://doi.org/doi:10.3949/ccjm.88a.20140
14. M. A. Marinescu, A. I. Löffler, M. Ouellette, L. Smith, C. M. Kramer, J. M. Bourque, Coronary microvascular dysfunction, microvascular angina, and treatment strategies, *JACC Cardiovasc. Imaging*, **8** (2015), 210−220. https://doi.org/10.1016/j.jcmg.2014.12.008
15. V. M. Pereyra, A. Seitz, H. Mahrholdt, R. Bekeredjian, U. Sechtem, P. Ong, Coronary microvascular dysfunction in patients with mild-to-moderate aortic stenosis - insights from intracoronary acetylcholine testing, *Int. J. Cardiol. Heart Vasc.*, **31** (2020), 100658. https://doi.org/10.1016/j.ijcha.2020.100658

16. E. Aribas, J. E. Roeters van Lennep, S. E. Elias-Smale, J. J. Piek, M. Roos, F. Ahmadizar, et al., Prevalence of microvascular angina among patients with stable symptoms in the absence of obstructive coronary artery disease: a systematic review, *Cardiovasc. Res.*, **118** (2022), 763−771. https://doi.org/10.1093/cvr/cvab061
17. *Branch of Cardiovascular Physicians*, *Chinese Medical Doctor Association*, 2020 expert consensus on the prevention and treatment of heart failure after myocardial infarction, *Chin. Circ. J.*, **35** (2020), 1166−1180. Available from: https://corestarbio.com/uploadfile/1669779085_2.pdf.
18. H. Liu, S. Ou, P. Liu, Y. Xu, Y. Gong, L. Xia, et al., Effect of microcirculatory resistance on coronary blood flow and instantaneous wave-free ratio: a computational study, *Comput. Methods Programs Biomed.*, **196** (2020), 105632. https://doi.org/10.1016/j.cmpb.2020.105632
19. J. Knuuti, W. Wijns, A. Saraste, D. Capodanno, E. Barbato, C. Funck-Brentano, et al., 2019 ESC Guidelines for the diagnosis and management of chronic coronary syndromes, *Eur. Heart J.*, **41** (2019), 407−477. https://doi.org/10.1093/eurheartj/ehz425
20. Y. Geng, X. Wu, H. Liu, D. Zheng, L. Xia, Index of microcirculatory resistance: state-of-the-art and potential applications in computational simulation of coronary artery disease, *J. Zhejiang Univ. Sci. B*, **23** (2022), 123−140. https://doi.org/10.1631/jzus.B2100425
21. P. G. Camici, G. d'Amati, O. Rimoldi, Coronary microvascular dysfunction: mechanisms and functional assessment, *Nat. Rev. Cardiol.*, **12** (2015), 48−62. https://doi.org/10.1038/nrcardio.2014.160
22. J. D. Sara, A. Sugrue, V. Kremen, B. Qiang, Y. Sapir, Z. I. Attia, et al., Electrocardiographic predictors of coronary microvascular dysfunction in patients with non-obstructive coronary artery disease: utility of a novel T wave analysis program, *Int. J. Cardiol.*, **203** (2016), 601−606. https://doi.org/10.1016/j.ijcard.2015.10.228
23. J. Herrmann, J. C. Kaski, A. Lerman, Coronary microvascular dysfunction in the clinical setting: from mystery to reality, *Eur. Heart J.*, **33** (2012), 2771−2783. https://doi.org/10.1093/eurheartj/ehs246
24. Z. Zhou, Y. Liu, Z. Wang, Y. Wang, J. Zhang, C. Yang, The value of standard deviation of normal RR Interval in predicting coronary microvascular dysfunction in patients with ischemia with non obstructive coronary arteries (in Chinese), *Chin. Circ. J.*, **37** (2022), 804−809. https://doi.org/10.3969/j.issn.1000-3614.2022.08.008
25. E. Fabris, A. W. J. van't Hof, 7- Angiography and electrocardiography (ECG) for the assessment of coronary microvascular obstruction, in *Coronary Microvascular Obstruction in Acute Myocardial Infarction* (eds. G. Niccoli and I. Eitel), New York: Academic Press, (2018), 109−126.
26. P. Ong, A. Athanasiadis, S. Hill, T. Schäufele, H. Mahrholdt, U. Sechtem, Coronary microvascular dysfunction assessed by intracoronary acetylcholine provocation testing is a frequent cause of ischemia and angina in patients with exercise-induced electrocardiographic changes and unobstructed coronary arteries, *Clin. Cardiol.*, **37** (2014), 462−467. https://doi.org/10.1002/clc.22282
27. N. Dose, M. M. Michelsen, N. D. Mygind, A. Pena, C. Ellervik, P. R. Hansen, et al., Ventricular repolarization alterations in women with angina pectoris and suspected coronary microvascular dysfunction, *J. Electrocardiol.*, **51** (2018), 15−20. https://doi.org/10.1016/j.jelectrocard.2017.08.017

28. J. D. Sara, R. J. Lennon, M. J. Ackerman, P. A. Friedman, P. A. Noseworthy, A. Lerman, Coronary microvascular dysfunction is associated with baseline QTc prolongation amongst patients with chest pain and non-obstructive coronary artery disease, *J. Electrocardiol.*, **49** (2016), 87−93. https://doi.org/10.1016/j.jelectrocard.2015.10.006
29. R. Alizadehsani, M. Abdar, M. Roshanzamir, A. Khosravi, P. M. Kebria, F. Khozeimeh, et al., Machine learning-based coronary artery disease diagnosis: a comprehensive review, *Comput. Biol. Med.*, **111** (2019), 103346. https://doi.org/10.1016/j.compbiomed.2019.103346
30. J. S. Richman, J. R. Moorman, Physiological time-series analysis using approximate entropy and sample entropy, *Am. J. Physiol. Heart Circ. Physiol.*, **278** (2000), H2039−H2049. https://doi.org/10.1152/ajpheart.2000.278.6.H2039
31. R. K. Udhayakumar, C. Karmakar, M. Palaniswami, Entropy profiling to detect ST change in heart rate variability signals, in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, (2019), 4588−4591. https://doi.org/10.1109/EMBC.2019.8857297
32. H. Rabbani, M. P. Mahjoob, E. Farahabadi, A. Farahabadi, A. M. Dehnavi, Ischemia detection by electrocardiogram in wavelet domain using entropy measure, *J. Res. Med. Sci.*, **16** (2011), 1473−1482. Available from: https://www.researchgate.net/publication/230843943.
33. J. Liu, C. Zhang, T. Ristaniemi, F. Cong, Detection of myocardial infarction from multi-lead ECG using dual-Q tunable Q-factor wavelet transform, in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, **2019** (2019), 1496−1499. https://doi.org/10.1109/EMBC.2019.8857775
34. C. Chen, Y. Jin, I. L. Lo, H. Zhao, B. Sun, Q. Zhao, et al., Complexity change in cardiovascular disease, *Int. J. Biol. Sci.*, **13** (2017), 1320−1328. https://doi.org/10.7150/ijbs.19462
35. J. M. Ahn, S. Arora, O. G. Solberg, O. Angerås, K. Rolid, M. Rafique, et al., Prognostic value of comprehensive intracoronary physiology assessment early after heart transplantation, *Eur. Heart J.*, **42** (2021), 4918−4929. https://doi.org/10.1093/eurheartj/ehab568
36. E. A. P. Alday, A. Gu, A. J. Shah, C. Robichaux, A. K. I. Wong, C. Liu, et al., Classification of 12-lead ECGs: the physioNet/computing in cardiology challenge 2020, *Physiol. Meas.*, **41** (2021), 124003−124003. https://doi.org/10.1088/1361-6579/abc960
37. X. Zhao, J. Zhang, Y. Gong, L. Xu, H. Liu, S. Wei, et al., Reliable detection of myocardial ischemia using machine learning based on temporal-spatial characteristics of electrocardiogram and vectorcardiogram, *Front. Physiol.*, **13** (2022). https://doi.org/10.3389/fphys.2022.854191
38. C. E. Kossmann, D. Brody, G. Burch, H. Hecht, F. Johnston, C. Kay, et al., Report of committee on electrocardiography, American Heart Association, Recommendations for standardization of leads and of specifications for instruments in electrocardiography and vectorcardiography, *Circulation*, **35** (1967), 583−602. https://doi.org/10.1161/01.cir.35.3.583
39. J. A. Kors, G. van Herpen, A. C. Sittig, J. H. van Bemmel, Reconstruction of the Frank vectorcardiogram from standard electrocardiographic leads: diagnostic comparison of different methods, *Eur. Heart J.*, **11** (1990), 1083−1092. https://doi.org/10.1093/oxfordjournals.eurheartj.a059647
40. J. Song, H. Yan, Z. Xiao, X. Yang, X. Zhang, A robust and efficientalgorithm for ST–T complex detection in electrocardiograms, *J. Mech. Med. Biol.*, **11** (2012), 1103−1111. https://doi.org/10.1142/s0219519411004198

41. A. Molina-Picó, D. Cuesta-Frau, M. Aboy, C. Crespo, P. Miró-Martínez, S. Oltra-Crespo, Comparative study of approximate entropy and sample entropy robustness to spikes, *Artif. Intell. Med.*, **53** (2011), 97−106. https://doi.org/10.1016/j.artmed.2011.06.007
42. M. A. Busa, R. E. A. van Emmerik, Multiscale entropy: a tool for understanding the complexity of postural control, *J. Sport Health Sci.*, **5** (2016), 44−51. https://doi.org/10.1016/j.jshs.2016.01.018
43. J. M. Yentes, N. Hunt, K. K. Schmid, J. P. Kaipust, D. McGrath, N. Stergiou, The appropriate use of approximate entropy and sample entropy with short data sets, *Ann. Biomed. Eng.*, **41** (2013), 349−365. https://doi.org/10.1007/s10439-012-0668-3
44. D. U. Uguz, F. Berief, S. Leonhardt, C. H. Antink, Classification of 12-lead ECGs using gradient boosting on features acquired with domain-specific and domain-agnostic methods, in *2020 Computing in Cardiology*, (2020), 1−4. https://doi.org/10.22489/CinC.2020.363
45. A. U. Haq, J. Li, M. H. Memon, M. H. Memon, J. Khan, S. M. Marium, Heart disease prediction system using model of machine learning and sequential backward selection algorithm for features selection, in *2019 IEEE 5th International Conference for Convergence in Technology (I2CT)*, 2019. https://doi.org/10.1109/I2CT45611.2019.9033683
46. M. Alnowami, F. Abolaban, E. Taha, A wrapper-based feature selection approach to investigate potential biomarkers for early detection of breast cancer, *J. Radiat. Res. Appl. Sci.*, **15** (2022), 104−110. https://doi.org/10.1016/j.jrras.2022.01.003
47. L. Li, X. Sui, J. Lian, F. Yu, Y. Zhou, Vehicle interaction behavior prediction with self-attention, *Sensors*, **22** (2022), 429. https://doi.org/10.3390/s22020429
48. C. Xie, R. Du, J. W. K. Ho, H. H. Pang, K. W. H. Chiu, E. Y. P. Lee, et al., Effect of machine learning re-sampling techniques for imbalanced datasets in 18F-FDG PET-based radiomics model on prognostication performance in cohorts of head and neck cancer patients, *Eur. J. Nucl. Med. Mol. Imaging*, **47** (2020), 2826−2835. https://doi.org/10.1007/s00259-020-04756-4
49. T. Saito, M. Rehmsmeier, The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets, *PLoS One*, **10** (2015), e0118432. https://doi.org/10.1371/journal.pone.0118432
50. S. Ansari, N. Farzaneh, M. Duda, K. Horan, H. B. Andersson, Z. D. Goldberger, et al., A review of automated methods for detection of myocardial ischemia and infarction using electrocardiogram and electronic health records, *IEEE Rev. Biomed. Eng.*, **10** (2017), 264−298. https://doi.org/10.1109/RBME.2017.2757953
51. R. Correa, P. D. Arini, L. S. Correa, M. Valentinuzzi, E. Laciar, Novel technique for ST-T interval characterization in patients with acute myocardial ischemia, *Comput. Biol. Med.*, **50** (2014), 49−55. https://doi.org/10.1016/j.compbiomed.2014.04.009
52. Y. L. Zang, L. Xia, Cellular mechanism of cardiac alternans: an unresolved chicken or egg problem, *J. Zhejiang Univ. Sci. B*, **15** (2014), 201−211. https://doi.org/10.1631/jzus.B1300177
53. M. J. Janse, A. L. Wit, Electrophysiological mechanisms of ventricular arrhythmias resulting from myocardial ischemia and infarction, *Physiol. Rev.*, **69** (1989), 1049−1169. https://doi.org/10.1152/physrev.1989.69.4.1049
54. C. C. ter Haar, A. C. Maan, S. G Warren, M. Ringborn, B. M. Horáček, M. J. Schalij, et al., Difference vectors to describe dynamics of the ST segment and the ventricular gradient in acute ischemia, *J. Electrocardiol.*, **46** (2013), 302−311. https://doi.org/10.1016/j.jelectrocard.2013.04.004

55. A. Feeny, L. G. Tereshchenko, Beat-to-beat determinants of the beat-to-beat temporal and spatial variability of repolarization, *J. Electrocardiol.*, **49** (2016), 417−422. https://doi.org/10.1016/j.jelectrocard.2016.01.007
56. M. Wei, J. Z. Song, H. Yan, Electrocardiogram ST-segment morphology variability analysis base on correlation coefficient entropy and inverse correlation coefficient entropy, *Appl. Mech. Mater.*, **195−196** (2012), 550−554. https://doi.org/10.4028/www.scientific.net/AMM.195-196.550
57. M. Ferrario, M. G. Signorini, G. Magenes, S. Cerutti, Comparison of entropy-based regularity estimators: application to the fetal heart rate signal for the identification of fetal distress, *IEEE Trans. Biomed. Eng.*, **53** (2006), 119−125. https://doi.org/10.1109/TBME.2005.859809
58. S. D. Fihn, J. C. Blankenship, K. P. Alexander, J. A. Bittl, J. G. Byrne, B. J. Fletcher, et al., 2014 ACC/AHA/AATS/PCNA/SCAI/STS Focused Update of the Guideline for the Diagnosis and Management of Patients With Stable Ischemic Heart Disease: A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines, and the American Association for Thoracic Surgery, Preventive Cardiovascular Nurses Association, Society for Cardiovascular Angiography and Interventions, and Society of Thoracic Surgeons, *J. Am. Coll. Cardiol.*, **64** (2014), 1929−1949. https://doi.org/10.1016/j.jacc.2014.07.017
59. H. A. Mayala, W. Yan, H. Jing, S. Liu, G. Yi, C. Qin, et al., Clinical characteristics and biomarkers of coronary microvascular dysfunction and obstructive coronary artery disease, *J. Int. Med. Res.*, **47** (2019), 6149−6159. https://doi.org/10.1177/0300060519859134

AIMS Press