ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 3 - 2026

# A heartbeat anomaly classification algorithm integrating dynamic parameter feature extraction method and random forest classification model

*Xiaochen Duan*[1], *Junbo Hao*[2], *Hankai Xu*[2], *Sicen Guo*[3] *and Ping Wang*[2,*]

[1] School of Big Data & Software Engineering,Chongqing University, Chongqing,400000, China
[2] School of Electrical Engineering, Chongqing University, Chongqing, 400000, China
[3] School of Microelectronics and Communication Engineering, Chongqing University, Chongqing, 400000, China

**SUMMARY:** *Conventional differential-threshold methods and machine-learning-based approaches often fail to guarantee both accuracy and real-time performance in heart rate variability (HRV) prediction. To address this issue, this paper proposes a heartbeat abnormality classification algorithm that integrates dynamic parameter feature extraction with a random forest classifier for fast and accurate electrocardiogram (ECG) analysis. First, raw ECG data are cleaned by digital filtering to obtain noise-suppressed signals, and dynamic thresholds are computed based on the heartbeats in the most recent 10 seconds. Second, a composite threshold score is calculated from the dynamic thresholds and current beat features to locate the QRS complex, followed by a local search for refinement. Finally, temporal feature vectors are constructed from the most recent heartbeats and fed into a pre-trained random forest model to obtain heartbeat abnormality classification results. Experimental results demonstrate that, compared with traditional differential-threshold algorithms and machine-learning-based methods, the proposed algorithm achieves both high accuracy and real-time performance in HRV prediction. On record 100 of the MIT-BIH Arrhythmia Database, the proposed method attains an accuracy of 99.65% and a sensitivity of 99.54%. Thus, relative to conventional differential-threshold and machine-learning-based approaches, the proposed algorithm maintains real-time capability while simultaneously achieving high classification accuracy.*

***KEYWORDS:*** *ECG detection; adaptive differential thresholding; band-pass filtering; wavelet transform; random forest; machine learning*

# 1 Foreword

Cardiovascular diseases (CVD) are the leading cause of death and disability worldwide. As a major risk factor, cardiac arrhythmia poses substantial challenges to clinical diagnosis and public health management. Electrocardiograms (ECGs), owing to their noninvasive nature and low cost, have been widely utilized for arrhythmia detection and monitoring [1]. With the rapid development of deep learning, automatic ECG analysis has achieved remarkable progress. In recent years, multi-task learning and uncertainty-aware methods have been introduced into ECG monitoring scenarios, improving model stability under complex conditions [1]. At the same time, self-supervised and continual learning strategies have enhanced model

*michael166@126.com
https://doi.org/10.65102/is20261220
Published: 10 June, 2026

generalization across datasets and application scenarios [2,3]. In addition, the emergence of lightweight neural architectures has pushed intelligent ECG analysis toward wearable devices and real-time monitoring [4]. Nevertheless, significant challenges remain in dynamic ECG monitoring. Existing models still struggle to reliably detect low-amplitude and subtle pathological patterns, such as T-wave alternans [5]; in settings with limited labeled data, deep models are prone to overfitting, which restricts their use in long-term monitoring [6]. More importantly, their robustness against interfering signals is inadequate, making them vulnerable to noise and abnormal inputs and thereby undermining clinical reliability [7].

In order to overcome these limitations, various improvement methods have been explored and studied. Activity-enhanced abnormal ECG signals synthesized by the Generative adversarial networks (GANs) were used to alleviate the problem of insufficient abnormal training samples [8]. However, significant deviations between the data generated by the GAN and the real ECGs are often found, which limits the model's generalization ability in the actual clinical environment, and the interpretability of the generation process itself is considered insufficient, which conflicts with clinical requirements for high credibility and traceability. A framework that combines attention mechanisms with self-encoders has been used to significantly enhance feature representation capabilities and arrhythmia classification performance [7], but such deep learning methods are still considered to rely heavily on large-scale annotated datasets, so performance in a small number of samples or real-time monitoring scenarios is considered not good enough. In addition, its black-box characteristics make the model's decision basis difficult for clinicians to understand, resulting in reduced clinical acceptance. At the same time, the combination of traditional signal processing techniques and deep learning models continues to be considered valuable. For example, arrhythmia detection based on high-resolution time-frequency analysis and convolutional neural networks can achieve high detection performance, but the sensitivity to signal preprocessing and parameter setting is considered too high to maintain stability under complex conditions of rapid changes in heart rate.[9]. Similarly, atrial fibrillation detection based on multi-lead signal time-adaptive deep networks improves complex waveform recognition capabilities, but problems with structural complexity and high computing cost have been discovered, hindering its deployment on resource-limited portable real-time monitoring devices [10].

All in all, although great progress in ECG intelligent analysis has been driven by existing methods, it is still plagued by obvious deficiencies in cross-scene adaptability, clinical interpretability, dynamic robustness and real-time computing efficiency. In order to make up for these shortcomings, an intelligent ECG analysis algorithm ADAA-RF that combines adaptive dynamic apodization and random forest classifier is proposed in this paper. During the signal preprocessing stage, an adaptive dynamic apodization mechanism is introduced. Through dynamic threshold and multi-dimensional feature extraction, the robustness of QRS wave detection under rapid fluctuations in heart rate and noise interference is enhanced; at the same time, the random forest classifier relies on the importance of features. Evaluation, model transparency and interpretability are improved, and more reliable decision support is provided to clinicians. Compared with deep learning methods such as convolutional neural networks, ADAA-RF maintains competitiveness in prediction accuracy while significantly reducing computational complexity, making it more suitable for portable devices and remote health monitoring scenarios. This framework not only fills an important gap in dynamic ECG feature modeling and clinical interpretability but also offers new theoretical support and engineering solutions for intelligent arrhythmia prediction [12].

# 2 Algorithm principle

The proposed heartbeat abnormality classification algorithm combines dynamic parameter feature extraction with a random forest classifier. Traditional digital filters are employed for data cleaning, while a hybrid scheme that couples the classical Pan–Tompkins algorithm with a random forest model is used for dynamic parameter feature extraction and rhythm discrimination. In addition, a regional search strategy is adopted to further improve detection accuracy. The overall algorithmic framework is illustrated in Figure 1.

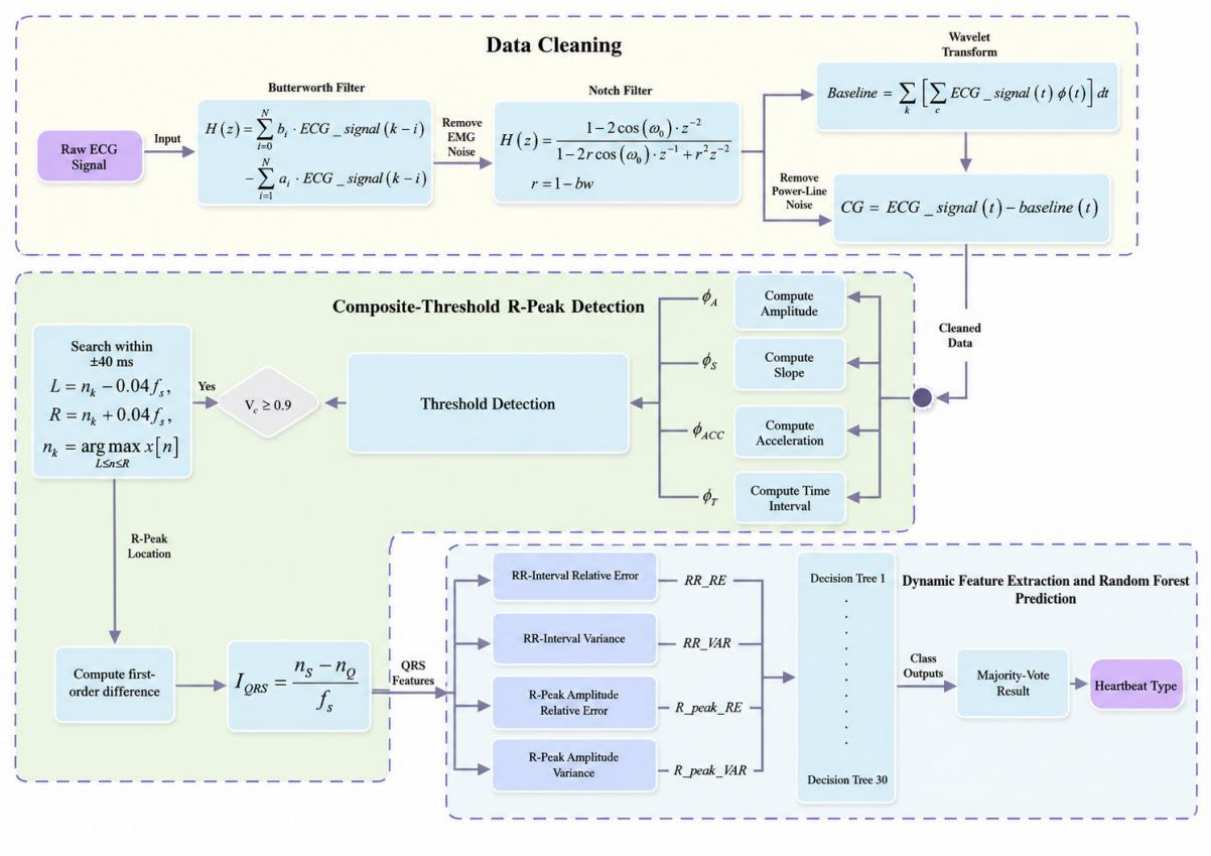


*Figure 1: Schematic diagram of the proposed algorithmic framework.*

## 2.1 Data Cleaning

Let the passband cutoff frequency $f_p$, stopband cutoff frequency $f_s$, maximum passband ripple $r_p$, and minimum stopband attenuation $r_s$ be specified. The order of the Butterworth filter $N$ can then be determined by the corresponding standard formula:

$$N = buttord(\omega_p, \omega_s, r_p, r_s) \tag{1}$$

among, $\omega_p = 2\pi f_p, \omega_s = 2\pi f_s$ represents the analog angular frequency of the passband and stopband.

Construct the zeros, poles, and gain coefficients of the normalized Butterworth analog filter, and transform the normalized filter into a low-pass filter with the desired cutoff frequency, expressed as:

$$H(s) = \frac{K_0 \prod_{i=1}^{M}(s - z_i)}{\prod_{j=1}^{N}(s - p_j)} \tag{2}$$

where $H(s)$denotes the frequency response function, $N_z$and $N_p$are the numbers of zeros and poles, respectively, $x(t)$is the input signal, $z_i$and $p_i$are the zeros and poles, and $K_0$ represents the normalized gain.

A bilinear transformation is then applied to convert the analog filter into a digital filter, which is subsequently used to filter the ECG signal $ECG\_signal(k)$:

$$H(z) = H(s)|_{s=\frac{2}{T}\cdot\frac{z-1}{z+1}}, y(k) = \sum_{i=0}^{N} b_i \cdot ECG_signal(k-i) - \sum_{i=1}^{N} a_i \cdot ECG_signal(k-i) \tag{3}$$

Here, $H(z)$denotes the complex frequency response of the filter, $z$ is the complex variable, $y(k)$ represents the output sequence, $a_i$ are the feedforward coefficients, and $b_i$ are the feedback coefficients.

An IIR notch filter is designed with a center frequency $f_0$ and quality factor $Q$. Its normalized center angular frequency is given by $\omega_0 = f_0/(f_s/2)$, and the bandwidth is $bw = \omega_0/Q$. This filter is used to suppress power-line interference in the ECG signal. The transfer function of the filter is expressed as:

$$H(z) = \frac{1 - 2cos(\omega_0) \cdot z^{-1} + z^{-2}}{1 - 2r\,cos(\,\omega_0) \cdot z^{-1} + r^2 z^{-2}}, r = 1 - bw \tag{4}$$

Wavelet decomposition is applied to the filtered electromyographic (EMG) signals and the power frequency (PFC) signals to eliminate baseline drift. The L-level wavelet decomposition can be expressed as:

$$c_{j,k} = \int_{-\infty}^{+\infty} ECG_signal(t) \cdot \psi_{j,k}(t)dt, j = 1,2,\dots,L \tag{5}$$

where $c_{j,k}$are the scaling coefficients, $\psi_{j,k}(t)$ denotes the Daubechies scaling function, and $L$ represents the set of wavelet basis functions.

By reconstructing the approximation components, the baseline wander component of the ECG can be obtained as:

$$Baseline(t) = \sum_{k} c_{L,k}\phi_{L,k}(t) \tag{6}$$

where $Baseline(t)$ denotes the baseline component and $\phi_{L,k}(t)$ is the scaling function. Subtracting the reconstructed baseline from the original signal yields:

$$ECG\_signal_{corrected}(t) = ECG\_signal(t) - Baseline(t) \tag{7}$$

which $ECG\_signal_{corrected}(t)$ is referred to as the cleaned ECG signal.

## 2.2 Dynamic Composite-Threshold R-Wave Detection

In this paper, a comprehensive threshold determination method is proposed, and four criteria are designed: amplitude threshold, gradient threshold, acceleration threshold and time threshold, which respectively reflect the characteristics of R peak amplitude, slope, acceleration and heart interval. The slope and acceleration can be expressed as:

$$s[n] = | x[n] - x[n-1] | \tag{8}$$

$$a[n] = | s[n] - s[n-1] | \tag{9}$$

where $s[n]$ denotes the slope of the R peak and $\mathrm{a}[n]$ denotes the corresponding acceleration.

The thresholds are adaptively updated in real time based on the heartbeats within the most recent 10 seconds. Specifically, the gradient threshold is set to 0.6 times the mean of the maximum slopes of the heartbeat waveforms in the last 10 seconds; the amplitude threshold is set to 0.5 times the mean of the maximum amplitudes; the acceleration threshold is set to 0.5 times the mean of the maximum accelerations (i.e., the second-order differences of the discrete sequence); and the temporal threshold is set to 0.3 times the mean RR interval of the most recent five beats. The composite threshold score is then defined as:

$$Vc = w_A\phi_A + w_S\phi_S + w_{Acc}\phi_{Acc} + w_t\phi_t, w_A + w_S + w_{Acc} + w_t = 1 \tag{10}$$

Here, $\phi_A$, $\phi_S$, $\phi_{Acc}$, and $\phi_t$denote the amplitude, slope, acceleration, and temporal scores, respectively. They are computed at each sampling instant to quantify how well the current point satisfies the R-peak criteria. The composite score $\mathrm{V_c} \in [0,1]$is obtained as a weighted sum of these four scores and reflects the overall likelihood that the current sample corresponds to an R peak. To balance robustness and sensitivity, the amplitude and slope terms are given greater weights, and all thresholds are dynamically updated over time through a sliding window approach.

The processed ECG signals are used to detect R-peak peaks by the proposed dynamic parameter differential composite threshold method. The algorithm computes the composite score $V_c$sample by sample on the continuously acquired data; when a sample satisfies $V_c \geq 0.9$, it is regarded as a likely R-peak candidate, thus completing the preliminary detection. To further ensure localization accuracy, once a candidate R-peak position is identified, the algorithm performs an exhaustive search within a $\pm 40$ msneighborhood, i.e., over the sample window $[n_k - 0.04f_s, n_k + 0.04f_s]$, to correct any R-peak localization bias. By conducting a local search on both sides of the candidate and locating the true maximum of the signal, the final R-peak position is aligned precisely with the actual R-wave apex of the ECG, and the resulting index $n_R$is taken as the final R-peak location:

$$L = n_k - 0.04f_s, R = n_k + 0.04f_s, n_R = argmax\, x[n] \tag{11}$$

The start and end boundaries of QRS wave group are extracted by local extremum combined with slope analysis, that is, the starting point of Q wave and the end point of S wave are determined. First, the first-order difference sequence of heart signal near R peak is calculated:

$$d[n] = x[n] - x[n-1] \tag{12}$$

Before the R peak, the algorithm searches for the Q-wave valley $n_{qv}$and the preceding local maximum $n_{qp}$. Within the interval $[n_{qp}, n_{qv}]$, the sample at which $| d[n] |$attains its

minimum (i.e., the smoothest transition point) is selected as the Q-wave onset $n_Q$. After the R peak, the S-wave valley $n_{sv}$and the subsequent local maximum $n_{sp}$are located; within the interval $[n_{sv}, n_{sp}]$, the sample with the minimum $|d[n]|$is taken as the S-wave endpoint $n_S$. Accordingly, the QRS time window can be expressed as:

$$T_{QRS} = \frac{n_S - n_Q}{f_s} \tag{13}$$

By jointly analyzing the minimum value of the derivative signal and the local extreme value of the original ECG signal, the starting point of the Q wave and the ending point of the S wave can be accurately located by the algorithm, thereby allowing the QRS complex to be completely outlined and ensuring the reliability of subsequent feature extraction [15]. If the estimated boundary exceeds the physiologically reasonable range, local backtracking search will be performed by the system to re-optimize the boundary, pseudo-extreme interference can be effectively suppressed, and stable segmentation performance is maintained.

## 2.3 Dynamic feature extraction algorithm

This study proposes an analysis framework that integrates dynamic feature extraction with a random forest classifier. Five types of dynamic characteristics are defined: QRS duration, relative error of heartbeat interval, variance of heartbeat interval, relative error of R-peak amplitude, and variance of R-peak amplitude. The proposed method is trained and evaluated on the MIT-BIH Arrhythmia Database, achieving high-accuracy recognition and classification of heartbeat abnormalities.

The formula for the QRS duration is:

$$Q_S = S_{end} - Q_{onset} \tag{14}$$

where $Q_S$denotes the duration of a single heartbeat, $S_{end}$ is the time of the S-wave end, and $S_{onset}$ is the time of the Q-wave onset.

The relative error of heartbeat interval is expressed as:

$$RR\_RE_n = \frac{\left|RR_n - \frac{1}{5}\sum_{i=1}^{5} RR_{n-i}\right|}{\frac{1}{5}\sum_{i=1}^{5} RR_{n-i}} \times 100\%, RR_n = R_{peak_n} - R_{peak_{n-1}} \tag{15}$$

where $RR_RE$denotes the relative error of the heartbeat interval, computed with respect to the previous five beats; $RR_n$represents the time interval of the $n$-th heartbeat, and $R_{\text{peak}_n}$is the time index of the R-wave peak in the $n$-th heartbeat signal.

The variance of the heartbeat interval is:

$$RR\_VAR_n = \frac{1}{6}\sum_{i=n-5}^{n}(RR_i - \frac{1}{6}\sum_{j=0}^{6} RR_{n-j})^2 \tag{16}$$

where $RR\_VAR$ denotes the variance of the heartbeat intervals (computed based on the differences between successive R-peak times), and the remaining symbols have the same meanings as in the formula for the relative error of the heartbeat interval.

The relative error of *R_peak* is expressed as:

$$R\_peak\_RE_n = \frac{\left| R\_peak_n - \frac{1}{6}\sum_{i=n-5}^{n} R\_peak_i \right|}{R\_peak_n} \times 100\% \tag{17}$$

where $R\_peak\_RE$ denotes the relative error of the R-peak amplitude, computed with respect to the previous five beats, and $R\_peak_n$ denotes the amplitude of the $n$-th R-wave peak.

The R peak variance expression is:

$$R\_peak\_VAR_n = \frac{1}{6}\sum_{i=n-5}^{n} (R\_peak_i - \frac{1}{6}\sum_{j=0}^{6} R\_peak_{n-j})^2 \tag{18}$$

where $\boldsymbol{R}_peak_VAR$denotes the variance of the R-peak amplitudes, and $\boldsymbol{R}_peak_n$denotes the amplitude of the $\boldsymbol{n}$-th R-wave peak.

The combination of the above dynamic statistical and morphological features enhances the robustness of discrimination across multiple rhythm types and provides more discriminative inputs for the subsequent classifier.

## 2.4 Random Forest Model

Random forest is a high-performance machine learning algorithm that can be used for both classification and regression. It reduces model variance and mitigates overfitting while maintaining high predictive accuracy and strong generalization capability. In the proposed method, random forests are employed to improve prediction accuracy without sacrificing computational efficiency [11, 12].

In this papers' algorithm, we assume the existing training set$D$:

$$D = \{(x_i, y_i)\}_{i=1}^{N} \tag{19}$$

where $\mathbf{x}_i \in \mathbb{R}^d$is the feature vector and $y_i \in \{1, \dots, C\}$is the corresponding label. The random forest consists of $n_d$decision trees $\{h_m(x)\}_{m=1}^{n_d}$.

In the proposed algorithm, considering the requirement for a lightweight model, we set $n_d = 100$, i.e., each random forest is composed of 100 decision trees. When training the $m$-th decision tree ($m \le n_d$), a bootstrap sample $D_m$is first drawn from the original training set, and the tree is then trained on this subsample. At each node, a random subset of features $F_m \subseteq \{1,2, \dots, d\}$is selected, and the split is chosen so as to maximize a predefined split-quality criterion.

For classification, the Gini function is commonly used in random forests to measure the quality of a split. It is defined as follows, where $p_i$denotes the probability of each class in the corresponding child node after the split:

$$Gini = 1 - \sum_{i=1}^{c} (p_i)^2 \tag{20}$$

At this point, the split criterion can be written as:

$$argmax_{f \in F_m, \theta} \Delta Gini(f, \theta) \tag{21}$$

where $\theta$denotes the split threshold.

In the proposed algorithm, the five parameter features described above are used as input features for training, and *Situation* is defined as the label representing the heartbeat type, i.e.

$$Situation = \begin{cases} 1, Normal\ heart\ beat \\ 0, Abnormal\ heart\ \text{beat} \end{cases} \tag{22}$$

In general, for a given input x, each decision tree outputs a class-probability distribution:

$$h_m(x) = \left(p_{m1}(x), \dots, p_{mC}(x)\right), \sum_{i=1}^{C} P_{mi}(x) = 1 \tag{23}$$

Finally, the classification result is obtained by averaging the probabilities over all decision trees:

$$\hat{y}(x) = argmax_{c\epsilon 1,\dots,C} \frac{1}{n_d} \sum_{i=1}^{n_d} P_{mc}(x) \tag{24}$$

By performing classification and prediction via the random forest, the proposed algorithm provides reliable decision support for subsequent ECG monitoring.

# 3 Experimental results analysis

## 3.1 Signal samples

In this work, ECG recordings from the MIT-BIH Arrhythmia Database are used as signal samples for analysis. The database contains 48 half-hour two-channel ambulatory ECG recordings from 47 subjects, with a resolution of 11 bits and a sampling frequency of 360 Hz.

## 3.2 Experimental Results

To validate the effectiveness of the proposed algorithm, we first consider record 100 of the database as an illustrative example. Figure 2 shows the first 3000 sampling points of this recording. As can be seen from Figure 2, the raw signal is heavily corrupted by noise, and a sequence of filtering operations is required before reliable R-peak detection can be performed.

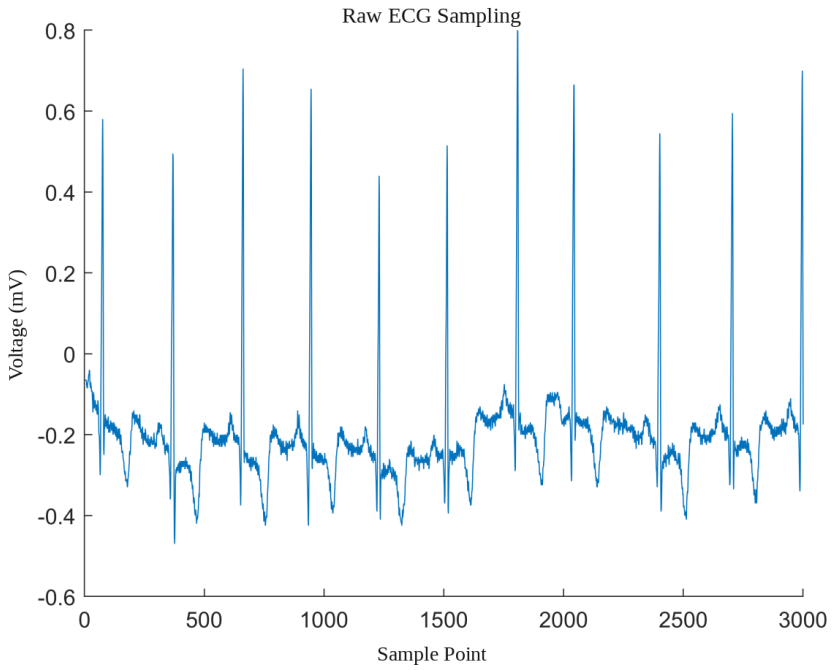


*Figure 2: Experimental data sampling diagram*

After removal of electromyographic noise, suppression of 50/60 Hz power-line interference, and elimination of baseline wander, the cleaned ECG samples are obtained as shown in Figure 3.

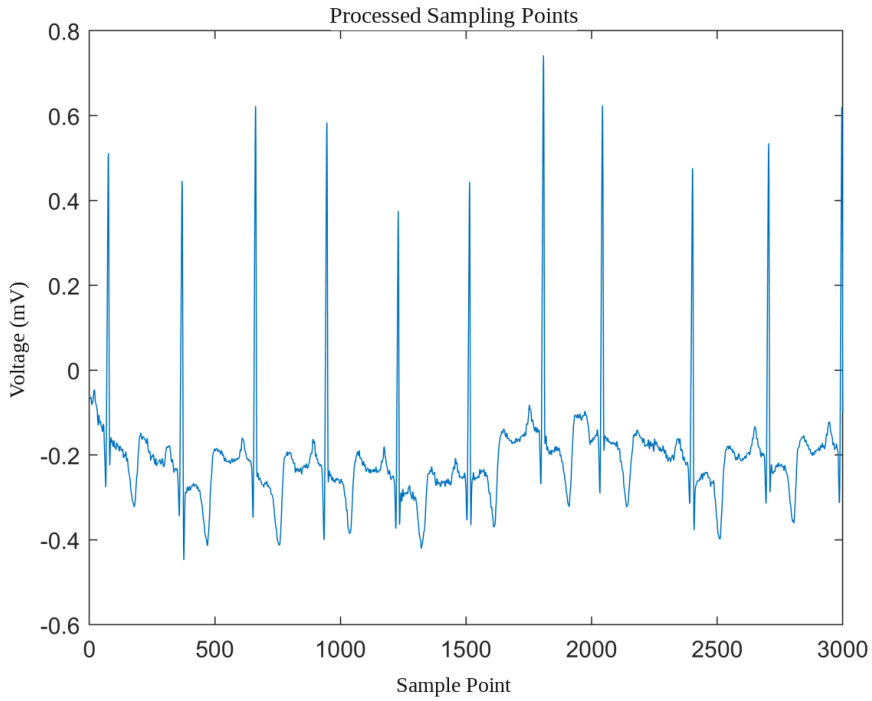


*Figure 3: Schematic of processed sampling points*

It is evident that noise has been significantly attenuated in the preprocessed signal.

The first minute of data from Record 100 (containing 21,600 sample points) is used for analysis to provide a complete assessment of system performance. The R-peak detection results obtained from these data are shown in Figure 4.

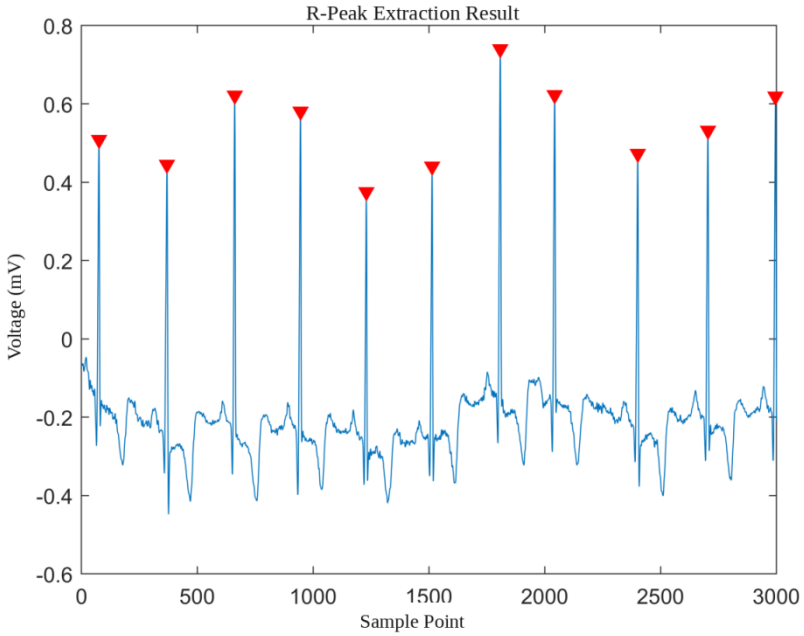


*Figure 4: Schematic diagram of R-wave peaks extracted by the algorithm*

As shown in Figure. 4, all R peaks in this ECG segment were successfully identified and marked with inverted triangles, which shows that the proposed algorithm has achieved good results in practical use.

## 3.3 Algorithm Effect Evaluation

TWhen the accuracy of the algorithm is quantitatively evaluated, the following performance indicators are adopted [13]:

Sensitivity ($Se$): The degree of missed detection is measured by this indicator; the proportion of real heartbeats that are successfully detected increases with the increase of sensitivity. It is defined as:

$$Se = \frac{TP}{TP+FN} \tag{25}$$

Accuracy ($ACC$): This metric measures the correctness of the detections; a higher accuracy implies that the detected heartbeats are more reliable. It is defined as:

$$ACC = \frac{TP}{TP+FP+FN} \tag{26}$$

Within a 50 ms search window, $TP$ denotes the number of R peaks correctly matched with the reference annotations, $FP$ denotes the number of falsely detected R peaks, and $FN$ denotes the number of missed R peaks.

Ten ECG records from the MIT-BIH database are selected and processed using the proposed algorithm. The performance of R-peak detection on these records is summarized in Table 1.

*Table 1: Detection of R-peak effect*

| Sample number | Actual heart rate | Check heart rate | Missed detection/ individual | Misidentified/1 | Sensitivity (%) | Precision (%) |
|---|---|---|---|---|---|---|
| 100 | 2274 | 2272 | 2 | 0 | 99.91 | 100 |
| 101 | 1874 | 1868 | 6 | 0 | 99.68 | 100 |
| 103 | 2091 | 2085 | 6 | 0 | 99.71 | 100 |
| 104 | 2311 | 2334 | 7 | 30 | 99.70 | 98.71 |
| 105 | 2691 | 2685 | 6 | 0 | 99.78 | 100 |
| 106 | 2098 | 2152 | 5 | 59 | 99.76 | 97.26 |
| 111 | 2133 | 2124 | 9 | 0 | 99.58 | 100 |
| 112 | 2559 | 2540 | 19 | 0 | 99.23 | 100 |
| 113 | 1796 | 1793 | 3 | 0 | 99.83 | 100 |
| 114 | 1890 | 1889 | 1 | 0 | 99.95 | 100 |
| amount to | 21717 | 21742 | 64 | 89 | 99.71 | 99.30 |

The selected 10 records cover a variety of heart rate conditions. For example, according to official documents, record 104 corresponds to a patient implanted with a cardiac pacemaker, while record 106 refers to severe ventricular ectopic beats. These conditions can introduce significant artifacts into the ECG waveform, which has a significant impact on detection accuracy and sensitivity. Experimental results show that the overall sensitivity of the proposed algorithm is 99.71%, and the accuracy rate is calculated to be 99.30%. These data are used to demonstrate that the method has strong anti-interference capabilities and high detection accuracy.

In addition, the proposed method is compared with the classic R-wave peak detection algorithm using the above indicators. By analyzing the ECG samples from the first minute of data point 100, the comparison results obtained are shown in Figure 5.

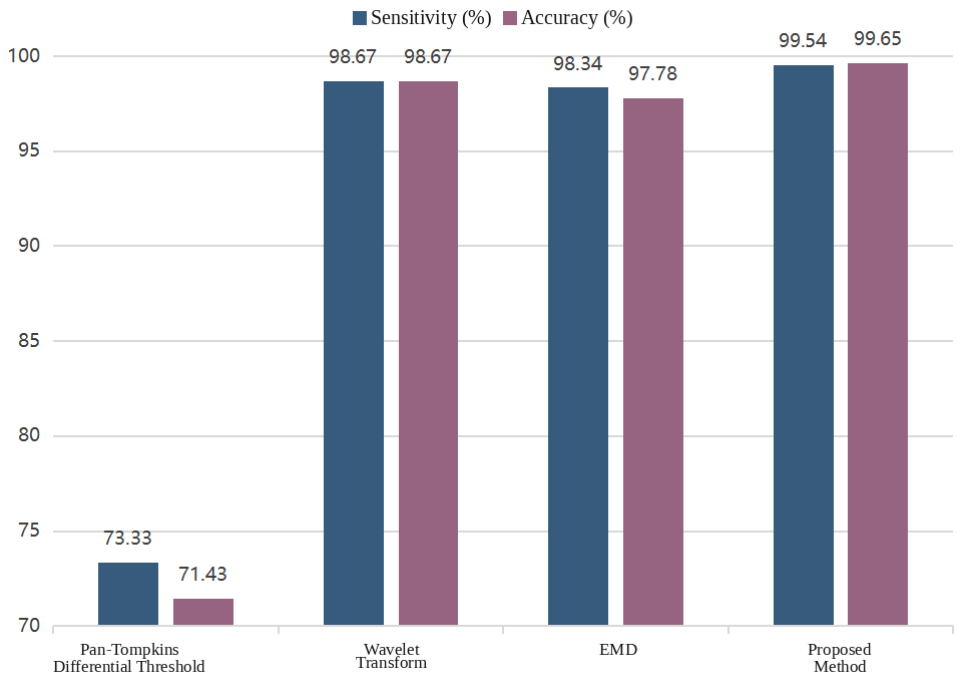


*Figure 5: Comparison of different algorithms*

The traditional Pan–Tompkins differential-threshold method is highly sensitive to noise and exhibits poor generalization to arrhythmic patterns, resulting in relatively low sensitivity and accuracy. Wavelet-transform-based methods are particularly responsive to the 5–25 Hz band, thereby highlighting QRS features and yielding higher accuracy; however, their algorithmic design is complex. Empirical mode decomposition (EMD) is based on empirical waveform analysis and lacks the strict mathematical foundation of the preceding methods; it also requires elaborate algorithmic design, which makes real-time analysis difficult, although it performs well for complex and rare rhythms. In contrast, the algorithm proposed in this paper offers high robustness, a simple and practical theoretical foundation, and superior accuracy.

## 3.4 Algorithm Performance Evaluation

Table 2 shows the time complexity of common heart rhythm detection algorithms, and Figure 6 shows the comparison diagram of the time complexity of the algorithms.

*Table 2: Time complexity of different algorithms*

| algorithm | time complexity |
|---|---|
| Pan-Tompkins Difference Threshold Method | $O(N \cdot win)$ |
| Wavelet Transform | $O(N)$ |
| Empirical Mode Decomposition (EMD) | $O(MS \cdot N + M \cdot N \cdot logN)$ |
| Neural network algorithm | $O(E \cdot N \cdot P)$ |
| Algorithm | $O(L \cdot S \cdot N \cdot D)$ |

Here, $win$ denotes the size of the sliding window; $M$ is the number of intrinsic mode functions (IMFs) obtained in EMD; $MS$ is a composite factor describing the interpolation complexity; $E$ denotes the number of training epochs for neural-network-based algorithms (set to 100 in this paper); $P$ is the total number of neural network parameters (approximately 3200 in our implementation); $L$ is the number of decision trees in the random forest (30 in this work); and $D$ is the number of features, which is 5 in our design.

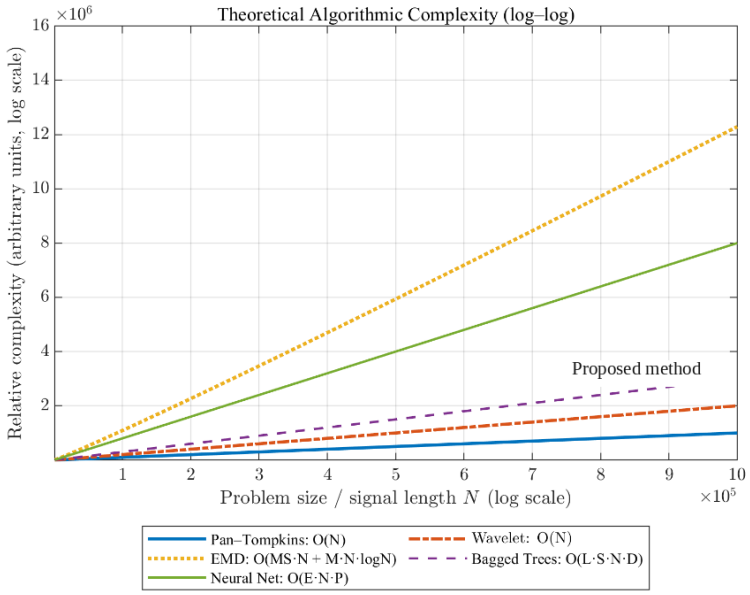


*Figure 6: Log-log comparison of algorithm complexities*

Current popular algorithms generally exhibit time complexity in linear or logarithmic order. While the Pan-Tompkins algorithm demonstrates high computational efficiency, its accuracy remains challenging to ensure. Although wavelet transform and EMD algorithms offer superior efficiency and accuracy, their algorithmic design typically relies heavily on empirical experience and mathematical validation, making them unsuitable for real-time applications. As machine learning approaches, neural networks and the proposed algorithm feature straightforward implementation, high efficiency, and accuracy. However, a large amount of training is required by neural network algorithms, which limits real-time performance. In contrast, fast decision tree training is achieved by the proposed algorithm (single tree complexity can reach a constant order), high accuracy is maintained in multiple decisions, low overfitting tendencies are demonstrated, and a retrospective monitoring mechanism is built into it, so significant algorithm advantages are demonstrated.

# 4 Conclusion

An ADAA-RF algorithm that combines dynamic parameter feature extraction and random forest classifier is proposed in this paper. By dynamically adjusting the decision threshold by adjacent heartbeat parameters, and adopting a threshold strategy that considers multiple dynamic features, composite time-series features are utilized in the random forest framework, so that the accuracy of heartbeat parameter identification and heartbeat abnormality classification can be simultaneously improved. Compared with the Pan-Tompkins differential-threshold method and wavelet transform method, the accuracy of ADAA-RF is improved by 39.51% and 2.01% respectively. Compared with methods based on neural-network-based and EMD-based methods, the time complexity of ADAA-RF is significantly reduced and the computing overhead is effectively reduced. After a small amount of computing efficiency is sacrificed, the classification accuracy and sensitivity are greatly improved, which makes the algorithm particularly suitable for deployment in portable monitoring terminals in practical applications.

# References

[1] Wang X, Gao H, Ma C, et al. Uncertainty-inspired multi-task learning in arbitrary scenarios of ECG monitoring[J]. IEEE Journal of Biomedical and Health Informatics, 2025, 29(6): 4046–4058.

[2] Aublin P G, Felblinger J, Oster J. A generalisable heartbeat classifier leveraging self-supervised learning for ECG analysis during magnetic resonance imaging[J]. IEEE Journal of Biomedical and Health Informatics, 2024, 28(9): 5147–5155.

[3] Gao H, Wang X, Chen Z, et al. ECG-CL: A comprehensive electrocardiogram interpretation method based on continual learning[J]. IEEE Journal of Biomedical and Health Informatics, 2023, 27(11): 5225–5236.

[4] Xiao J, Liu J, Yang H, et al. ULECGNet: An ultra-lightweight end-to-end ECG classification neural network[J]. IEEE Journal of Biomedical and Health Informatics, 2021, 26(1): 206–217.

[5] Plaza-Seco C, Baksh M, Barner K E, et al. DeepTWA-TM: Deep learning T-wave alternans detection in ambulatory ECG via time analysis[J]. IEEE Journal of Biomedical

and Health Informatics, 2025: in press.

[6] Ma C, Xiao Z, Zhao L, et al. A review on atrial fibrillation detection from ambulatory ECG[J]. IEEE Transactions on Biomedical Engineering, 2023, 71(3): 876–892.

[7] Han X, Hu Y, Foschini L, et al. Deep learning models for electrocardiograms are susceptible to adversarial attack[J]. Nature Medicine, 2020, 26(3): 360–363.

[8] Chen L, Xiao S, Zang Y, et al. GAEA-Net: Generating activity-enriched abnormal ECGs via adversarial network[J]. IEEE Journal of Biomedical and Health Informatics, 2025: in press.

[9] Tripathi P M, Kumar A, Kumar M, et al. Multilevel classification and detection of cardiac arrhythmias with high-resolution superlet transform and deep convolution neural network[J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 1–13.

[10] Zhang X, Jiang M, Polat K, et al. Detection of atrial fibrillation from variable-duration ECG signal based on time-adaptive densely network and feature enhancement strategy[J]. IEEE Journal of Biomedical and Health Informatics, 2022, 27(2): 944–955.

[11] Breiman L. Random forests[J]. Machine Learning, 2001, 45: 5–32.

[12] Zou C, et al. Heartbeat classification by random forest with a novel context feature: A segment label[J]. IEEE Journal of Translational Engineering in Health and Medicine, 2022, 10: 1–8.

[13] Baratloo A, Hosseini M, Negida A, et al. Part 1: Simple definition and calculation of accuracy, sensitivity and specificity[J]. Emergency (Tehran), 2015, 3(2): 48–49.

[14] Friesen G M, Jannett T C, Jadallah M A, et al. A comparison of the noise sensitivity of nine QRS detection algorithms[J]. IEEE Transactions on Biomedical Engineering, 1990, 37(1): 85–98.

[15] Khan W A, Rehman B, Abid M I, et al. A continuous wavelet and fast Fourier transform-based single-phase adaptive auto-reclosing scheme for EHV transmission lines[J]. Electrical Engineering, 2023, 105(3): 1347–1361.

[16] Petrović P, Damljanović N. New procedure for harmonics estimation based on Hilbert transformation[J]. Electrical Engineering, 2017, 99: 313–323.

[17] Diamandi I, Sahalos J N. Frequency and time domain analysis of the scattered field of buried dielectric targets[J]. Archiv für Elektrotechnik, 1994, 77: 441–449.

[18] Hannun A Y, Rajpurkar P, Haghpanahi M, et al. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network[J]. Nature Medicine, 2019, 25(1): 65–69.

[19] Attia Z I, Noseworthy P A, Lopez-Jimenez F, et al. An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction[J]. The Lancet, 2019, 394(10201): 861–867.

[20] Gadaleta M, et al. Prediction of atrial fibrillation from at-home single-lead ECG signals without arrhythmias[J]. npj Digital Medicine, 2023, 6(1): 229.