ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

UDC 519.7:65.011

doi: 10.32620/reks.2026.1.02

Mykola YEFREMOV[1], Andrey LIASHKO[1], Iurii KRAK[1,2], Oleg STELIA[1,2]

[1] *Taras Shevchenko National University of Kyiv, Kyiv, Ukraine*
[2] *Glushkov Institute of Cybernetics, Kyiv, Ukraine*

AN EFFICIENT THRESHOLD METHOD FOR DETECTING R-PEAKS IN ELECTROCARDIOGRAM

Automated R-peak detection in electrocardiogram (ECG) signals is essential to heart rhythm analysis, with applications in heart rate monitoring, heart rate variability (HRV) assessment, arrhythmia diagnosis ets. Its accuracy, however, is highly sensitive to noise and artifacts present in the ECG recording. The study proposes a method and its software implementation for R-peak detection, based on signal smoothing followed by differentiation and the application of a thresholding approach. ***The method*** *is designed for use in resource-constrained environments, such as portable and embedded monitoring systems.* ***The objective*** *of the study is to develop a computationally efficient and accurate method for the automatic detection of R-peaks in ECG signals, tailored to a personalized patient approach. The primary focus is on robustness to noise and QRS complex morphology, as well as the algorithm's ability to operate under limited computational resources and in real-time conditions. The proposed method relies on staged ECG signal processing. To validate the approach and compare its effectiveness, existing studies in the field were analysed.* ***The subject*** *of the study involves processing ECG signals and addressing challenges in R-peak detection in ECGs recorded via Holter monitors by constructing a smoothed continuous signal based on discrete data obtained during digitization. The discrete nature of the initial signal complicates differentiation and the precise identification of characteristic points, specifically R-peaks, which play a crucial role in diagnosing cardiac conditions.* ***The scientific novelty*** *of this investigation lies in the use of a second-order piecewise polynomial approximation to represent the ECG signal. This approach enables noise reduction in the signal and represents the discrete signal as a continuous function together with its first derivative, thereby permitting the analytical computation of its derivative.* ***Results*** *involve the detection of R-peaks by analysing the derivative of the smoothed signal: regions with sharp changes characteristic of the QRS complex were identified, and an iterative smoothing scheme was developed, with the number of iterations determined by a proposed stopping criterion. The proposed method was implemented in software and tested on data from the open-access MIT-BIH Arrhythmia Database, which includes over 60 recordings and more than 100,000 annotated R-peaks. The results were compared with studies by other authors using the F_1 score metric, based on standard precision and sensitivity metrics.* ***Conclusions:*** *The study proposes an effective and adaptive approach to ECG signal processing, ensuring reliable R-peak detection under conditions of significant noise, baseline drift, and physiological variability across patients. The obtained results demonstrated high performance metrics: up to 99.5% (average above 99.1%), F_1 score consistently exceeding 99%, and in some recordings reaching 100%. Thus, the proposed approach is competitive, demonstrating high accuracy in detecting R-peaks in cases of arrhythmias, peak inversions, non-standard QRS complex morphology, and other challenging signal conditions. The study's results can be applied in ECG analysis practice, particularly in the development of automated diagnostic systems or signal preprocessing before the application of classification methods.*

Keywords: *electrocardiogram; piecewise-polynomial approximation; R-peaks; smoothing; differentiation; MIT-BIH.*

1. Introduction

In modern society, the integration of digital technologies into healthcare is rapidly increasing, opening new opportunities for enhancing diagnostic accuracy and advancing the automation of biomedical signal processing. Electrocardiography is a vital tool for diagnosing and monitoring heart conditions, capturing the heart's electrical activity throughout each cardiac cycle. A normal electrocardiogram (ECG) consists of the P wave, the QRS complex and the T wave, as illustrated in Fig.1. These components are key for identifying normal and abnormal cardiac rhythms.

Automated ECG analysis is widely used in clinical practice to monitor heart rate, assess heart rate variability (HRV), and identify arrhythmias and other cardiac abnormalities. However, reliable R-peak detection remains challenging in the presence of noise, requiring effective denoising techniques that preserve the signal's essential characteristics, robust feature extraction, and precise localization of R-peaks.

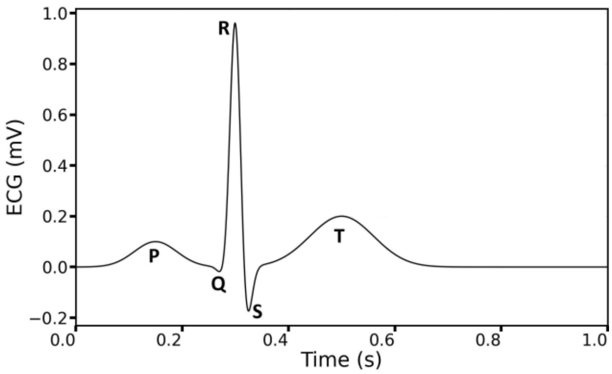


Fig. 1. A single cardiac cycle for a typical ECG

The presented research focuses on the development of a threshold-based algorithm for R-peak detection, based on differentiation of a smoothed signal. A prominent method in this category is the Pan-Tompkins approach (J. Pan, W. J. Tompkins) [1], along with its subsequent improvements [2]. This method facilitates adaptation to varying noise levels, automatic detection of anomalous signal segments, and maintains high accuracy in complex scenarios, making it promising for clinical practice and real-time monitoring systems.

1.1. Motivation

Automated analysis of ECG patterns is an essential element of modern cardiovascular diagnostics, as it alleviates the workload of healthcare professionals, enhances the precision of identification of pathological conditions, and promotes unbiased ECG signal analysis. A critical step in this process is R-peak detection within QRS complexes, since these cardinal reference points serve as fiducial benchmarks for heart rate computation and for deriving characteristic parameters such as RR, PR, and QT intervals, which are subsequently utilized to facilitate precise clinical diagnosis.

The accuracy of R-peak detection is critical for the reliability of subsequent stages of ECG analysis, including the diagnosis of arrhythmias, ischemic changes, and hypertrophic processes. Automatic R-peak detection is especially important for mobile monitoring, such as with Holter monitors or portable wearable devices used outside clinical environments. These systems face strict limitations in computational power, power consumption, and memory capacity, requiring algorithms with low computational complexity and minimal memory footprint. In real-world conditions, ECG signals are subject to various types of noise, including powerline interference, muscle contractions, motion artifacts, baseline wander, and variations in the physiological state of the individual, which affect the morphology of the QRS complex. These factors significantly complicate the reliable detection of R-peaks.

It is noteworthy that the advancement of medical practice, which tailors diagnostic and therapeutic strategies to the specific characteristics of each patient, imposes distinct requirements on the instruments used for acquiring ECG signals, the hardware employed for data transmission and storage, and the mathematical methods and algorithms utilized for subsequent ECG analysis to enhance diagnostic precision and convenience for medical specialists.

Devices designed for ECG signal acquisition are increasingly compact and straightforward to operate, comparable in simplicity to a standard wristwatch. Wearable technologies facilitate automated data transmission, providing alerts to users regarding potential irregularities in cardiac function. The transmitted ECG data undergo several stages of processing through sophisticated mathematical algorithms, which extract characteristic features of the signal, identify potential artifacts, and highlight segments indicative of possible cardiac abnormalities. Nevertheless, ECG recordings obtained in ambulatory settings are vulnerable to interference from various sources of high-frequency and low-frequency noise.

Consequently, there is an increased need for advanced mathematical methods and algorithms for ECG signal processing and analysis. The development of low-complexity algorithms, particularly those capable of achieving high accuracy in detecting R-peaks in the presence of significant signal noise, is therefore essential.

1.2. State of the art

In recent years, researchers have actively studied various methodologies for improving the accuracy of R-wave detection in ECG signals, taking into account the presence of noise and variations in data quality.

The review paper [3] presents a detailed analysis of computer methods proposed for R-peak detection in recent years, in particular, seven approaches are highlighted, namely: Wavelet Transform, Digital Filtering, Mathematical Morphology, Differentiator, Shannon Energy and Hilbert, Transform, Hybrid, Machine Learning. An earlier review paper [4] also explores similar approaches to R-peak detection: Pan and Tompkins method, Wavelet Transform, Empirical Mode Decomposition, Hilbert-Huang Transform, Fuzzy logic systems, Artificial neural networks. The review paper [5] focuses on the combination of wavelet transform algorithms and neural networks. Therefore, the use of wavelet transforms is one of the most common techniques.

It should be noted that the automation of the process of analyzing ECG signals based on wavelet transform [6] revealed significant difficulties for using this method in real time and for implementation on mobile (wearable)

ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

devices with limited computing power.

It is for important modern applications that it is necessary to develop other mathematical tools, in particular, threshold methods based on filtering the ECG signal and its subsequent differentiation, which are fully adapted to solve such problems. Therefore, let us dwell in more detail on the works devoted to threshold methods for determining R-waves [1,2].

In work [7], the authors apply adaptive threshold methods for accurate localization of R-waves as dominant elements of the signal. In the proposed development, R-waves are first determined, and then Q, S, T and P waves and their time characteristics are identified. In [8], an approach based on Slope Change Coefficients, obtained by integrating two scaled signals with opposite signs, is proposed. Although this method is not exclusively thresholding, it is combined with thresholding techniques for R-wave segmentation, taking into account slope changes as an additional criterion, which increases the robustness to noise and signal variations. In [9], the Hilbert transform is used to detect QRS complexes in combination with the dynamic energy selection (DMSE) and adaptive window size (AWS) algorithms. The DMSE algorithm effectively separates QRS components from unwanted signal components, and the adaptive window allows for dynamic adjustment of thresholds depending on the signal characteristics, ensuring high accuracy in variable conditions.

In [10], the Pan-Tompkins algorithm [1, 2] is used to detect QRS complexes in the context of diagnosing cardiovascular abnormalities for ambulatory monitoring. This algorithm is characterized by its simplicity of implementation and low computational requirements, which makes it effective for portable systems. However, as noted in [11], this approach has limitations when processing noisy ECG signals. The authors emphasize the need to improve this algorithm or develop new methods to increase the reliability of QRS complex detection in difficult conditions.

In [12], a combination of the Penn-Tompkins algorithm with the Hilbert transform is proposed, which calculates an analytical signal for analyzing the instantaneous amplitude and phase, which contributes to a more accurate identification of rhythm disturbances. In [13], a modified Pan-Tompkins algorithm is presented for accurate determination of R-waves. The problem of inaccurate R-wave annotation [14] in the MIT-BIH Arrhythmia database [15] from PhysioNet [16] is addressed.

A review [17] examines methods for validating R-wave detection algorithms, including thresholding approaches, before their implementation in real time. The authors emphasize the importance of transparent testing procedures to avoid false positives and ensure objective comparisons. Particular attention is paid to the challenges associated with noise and inaccurate annotations in databases such as MIT-BIH Arrhythmia [16] and the need for standardized metrics for evaluating thresholding methods.

Thus, the thresholding methods considered in these works demonstrate significant progress in improving accuracy and noise robustness through adaptive thresholding, signal preprocessing, and combination with other techniques such as slope analysis or Hilbert transforms. The algorithms described in [9] and [13] are characterized by their simplicity of implementation and efficiency in systems with limited resources. However, as noted in many studies, R-wave detection methods require further improvement, in particular in optimizing the search window size (150–300 ms, which corresponds to standard protocols) and increasing accuracy in noisy conditions. This remains an important area for future research.

ECG recording devices are highly sensitive devices, as a result of which they record various noises arising from sources such as interference in power supply lines, motion artifacts due to cable movement or electrode placement problems, as well as ECG interference due to other electrical activity. Noise suppression of the ECG signal is an integral part of the computer signal processing algorithm. Such pre-processing helps to increase the accuracy of extracting useful information from the signal. In [18], an unbiased finite impulse response filter (UFIR [19]) is used to determine the duration and amplitude of the P wave, QRS complex, and T wave on a standard ECG signal map.

One of the frequently used noise removal methods is Savitsky-Goley (SG) filtering, a classical signal smoothing method based on local least-squares approximation of the analyzed signal [20]. A review and comparative analysis of filters for filtering ECG signals are given in [21].

Thus, the analysis of ECG signal processing methods and the identification of their important characteristics from the point of view of applications can be divided into several groups: (1) – classical differentiation and thresholding methods. Advantages: speed. Disadvantages: sensitivity to noise; (2) – time-frequency analysis methods. Advantages: better feature localization. Disadvantages: computational complexity for embedded systems; (3) – methods based on artificial intelligence and machine learning. Advantages: high accuracy. Disadvantages: need for GPU/powerful processors, the “black box” problem; (4) – approximation and regression methods. Advantages: combination of analytical accuracy (due to the continuity of derivatives) with low hardware requirements. Disadvantages: use of ECG signal smoothing.

The presented article proposes an approach to implementing a threshold method that allows for accurate detection of R-waves in the ECG signal. Compared with existing literature reviews [22], the proposed method is based on a piecewise polynomial representation of the ECG signal.

The piecewise polynomial approximation facilitates attaining the desired level of signal smoothness based on the intensity of the noise. The approximation transforms the discrete signal into a continuous function, along with its first derivative, which can be analytically differentiated. The proposed algorithm requires minimal computational resources due to the use of unified expressions for both signal smoothing and subsequent differentiation. The algorithm demonstrates effective performance even in the presence of significant artifacts, making it suitable for ambulatory monitoring applications using wearable devices during patient movement.

The proposed approach facilitates adaptive signal processing tailored to the individual characteristics of the patient, reduces the impact of noise, and ensures high accuracy in detecting R-peaks across diverse clinical scenarios.

1.3. Objectives and tasks

The objective of this study is to develop and investigate an efficient threshold-based method for detecting R-peaks in ECG signals, utilizing a piecewise polynomial representation of the signal. To achieve the stated objective, the following research tasks have been defined:

1. Develop an algorithm for smoothing discrete ECG signals that effectively suppresses noise while preserving diagnostically significant features.
2. Develop and apply the steps of the algorithm for detecting R-peaks.
3. Establish criteria for terminating the iterative smoothing process to balance noise suppression efficiency with the preservation of signal informativeness.
4. Evaluate the accuracy and reliability of the algorithm using metrics such as Recall (Se), Precision (PPV) and F_1 score for R-peak detection on ECG data from the publicly available MIT-BIH Arrhythmia Database [15].

The objective is to develop an algorithm that achieves:

– F_1 score ≥ 99.1% as an average across the full MIT-BIH database;

– Recall Se > 99.5% for records with standard QRS morphology;

– Computational efficiency allowing real-time processing on 8-bit or 32-bit microcontrollers without an FPU (Floating Point Unit), specifically by reducing the operation count.

The paper has the following structure. Section 2 provides a detailed account of the materials and methods employed. This includes a thorough explanation of the piecewise polynomial approximation method (2.1), the step-by-step R-peak detection algorithm (2.2), the critical termination condition for the smoothing process (2.3), and the metrics used for accuracy evaluation (2.4).

Section 3 presents the experimental results, detailing the dataset used (3.1) and the outcomes (3.2). Section 4 offers a detailed discussion, analysing the algorithm's performance, particularly on challenging ECG records, and interpreting the significance of its behaviour in cases of false positives or missed detections.

Finally, Section 5 concludes the paper by summarizing the key contributions, affirming the high efficiency and adaptability of the proposed approach, and highlighting its potential clinical relevance in correctly classifying medically important, though atypical, waveforms as "suspicious" for further physician review.

2. Materials and methods of research

During the digitization of the ECG signal, the analog signal is converted into a sequence of discrete function values, which imposes limitations on its subsequent processing. Specifically, the discrete nature of the signal complicates the application of methods based on differentiation. To achieve more accurate localization of R-peaks and to adapt the method to varying recording conditions and individual patient characteristics, a piecewise polynomial approximation is employed [22].

2.1 The method of piecewise polynomial approximation of ECG signal data

A piecewise polynomial approximation is developed for the grid function, employing a Hermite polynomial with multiple nodes, designed for the identification of R-peaks in ECG signals.

Consider a sequence of points τ_i, distributed over the interval $[\tau_1, \tau_N]$.

For each value τ_i, a corresponding value of the grid function F_i is defined. Over the specified interval, the grid function is approximated by a piecewise polynomial function.

For the piecewise polynomial approximation of the continuous function together with its first derivative, an auxiliary grid with nodes x_i is introduced:

$$\tau_1 \le x_2 <\ldots< x_N \le \tau_N,$$

where

$$x_i = (\tau_{i-1} + \tau_i)/2,\ \ i = \underline{2, N}.$$

For each of the intervals $[x_i,\ x_{i+1}]$, where

ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

$i = \underline{2, N-1}$ the corresponding Hermite polynomials are constructed. The function values f_i and their derivatives f_i' are defined at the points x_i as follows:

$$f_i = (F_i + F_{i-1})/2, \quad f'_i = \frac{F_i - F_{i-1}}{\tau_i - \tau_{i-1}}, \; i = \underline{2, N}.$$

The polynomials are then written for each of the intervals $[x_i, x_{i+1}]$:

$$H_{3i}(x) = f_i \frac{x - x_{i+1}}{x_i - x_{i+1}} + f_{i+1} \frac{x - x_i}{x_{i+1} - x_i} + \\ +(x - x_i)(x - x_{i+1})(a_i x + b_i).$$

For the polynomials $H_{3i}(x)$, the following conditions are imposed:

$$H_{3i}(x_i) = f_i, H_{3i}(x_{i+1}) = \\ = f_{i+1}, H'_{3i}(x_i) = f'_i, H'_{3i}(x_{i+1}) = f'_{i+1}.$$

The unknown parameters are determined from the system:

$$f'_i = f_i \frac{1}{x_i - x_{i+1}} + f_{i+1} \frac{1}{x_{i+1} - x_i} + \\ +(x_i - x_{i+1})(a_i x_i + b_i),$$

$$f'_{i+1} = f_i \frac{1}{x_i - x_{i+1}} + f_{i+1} \frac{1}{x_{i+1} - x_i} + \\ +(x_{i+1} - x_i)(a_i x_{i+1} + b_i).$$

For a uniform arrangement of nodes with step h, the following expressions are obtained

$$a_i = 0,$$

$$b_i = \frac{f'_{i+1} - f'_i}{2h} - \frac{(f'_{i+1} + f'_i)(x_i + x_{i+1})}{2h^2} + \\ + \frac{(f_{i+1} - f_i)(x_i + x_{i+1})}{h^3}.$$

It should be noted that the constructed polynomial is quadratic.

The desired polynomial is expressed as follows:

$$H_{2i}(x) = f_{i+1} \frac{x - x_i}{h} - f_i \frac{x - x_{i+1}}{h} + \\ +(x - x_i)(x - x_{i+1})b_i.$$

The polynomials $H_{2i}(x)$ form a function that is continuous together with its first derivative. Hence, the derivative is expressed as:

$$H'_{2i}(x) = \frac{F_{i+1} - F_{i-1}}{2h} + (2x - x_i - x_{i+1})b_i.$$

2.2 Algorithm for R-peak detection based on piecewise polynomial approximation

The algorithm for detecting R-peaks consists of the following series of steps:

1. **Signal smoothing:** Although modern Holter monitors have built-in filters, the high sensitivity of such devices does not allow for obtaining a perfect signal for further analysis, and numerical differentiation of a discrete function increases noise. Therefore, before detecting R-peaks, the proposed algorithm performs preliminary smoothing of the ECG signal.

In practice, for an ECG that contains low-level noise, 2–3 steps of sequential smoothing are usually sufficient. In the presence of artifacts and significant noise, the number of iterations increases, but when increasing the number of iterations, it is important not to lose the informational content of the signal. Typically, for a heavily noisy cardiogram, 150 iterations are sufficient. If the first iteration of constructing the smoothing polynomial does not yield a satisfactory result, further iterations are performed. In cases where the selected ECG fragment contains minimal noise, a single smoothing iteration is sufficient. The smoothed signal is used for further analysis. Fig. 2 shows the results of signal smoothing under varying noise intensities.

2. **Differentiation of the smoothed signal:** The values of the derivatives are calculated at points where the signal is a decreasing function. This selection of points for derivative computation is motivated by the fact that the right slope of the QRS complex is typically steeper.

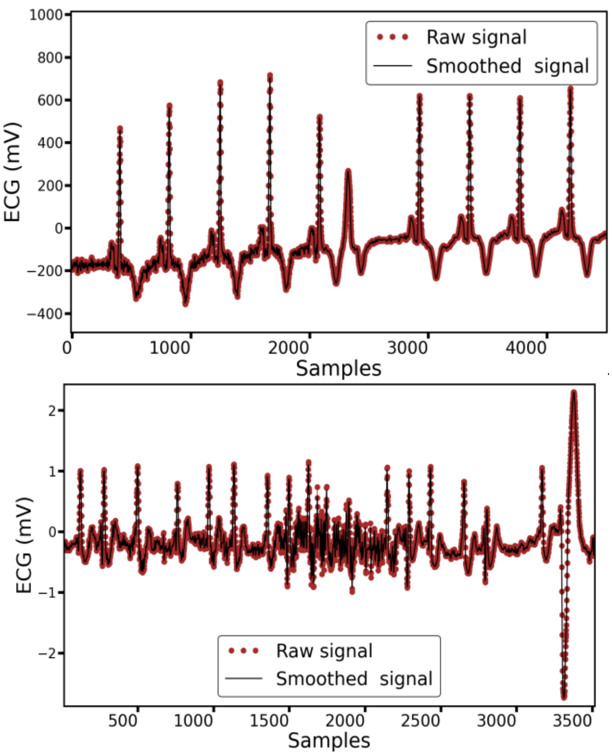
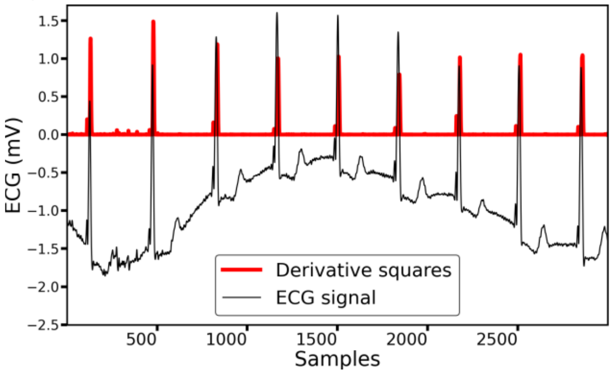


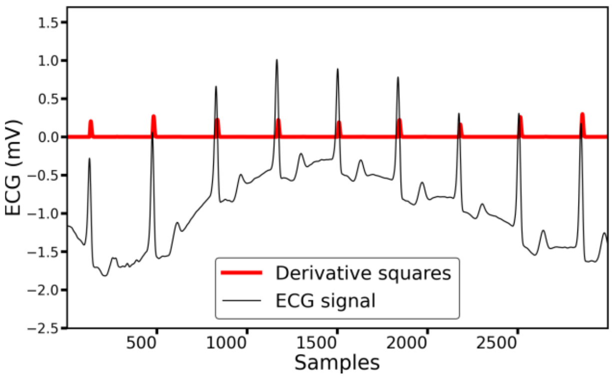
Fig. 2. Examples of smoothed ECG signal fragments

At all other points, the derivative value is set to zero. This approach reduces the number of points for R-peak detection by nearly half. The derivative values are then squared. Squaring the derivative values facilitates the localization of potential R-peak positions. Figure 3 illustrates an ECG segment after smoothing, along with the computed squares of the derivatives. Increasing the number of smoothing iterations for the initial signal enhances the prominence of R-peaks, thereby improving analysis and further processing.

As shown in Fig. 3a and Fig. 3b, increasing the number of smoothing iterations from 3 to 150 substantially enhances the prominence of squared derivative values at probable R-peak locations, making them clearly distinguishable from surrounding signal regions.



(a)



(b)

Fig. 3. ECG Signal with varying smoothing intensities and corresponding squares of the first derivative: (a) ECG segment after three smoothing iterations; (b) ECG segment after 150 smoothing iterations

3. Detection of intervals containing localized R-peaks: The threshold value is determined based on the computed squares of the derivatives. For each patient, the threshold is individually tailored and may serve as an additional characteristic of the patient's ECG. Subsequently, intervals are identified that encompass points where the squared derivatives exceed or equal the threshold. The number of points within each such interval is chosen to not exceed the point count associated with the QRS complex on the specified cardiograph.

4. R-peak detection: At the final step of the algorithm, the intervals identified in the previous step are sequentially processed.

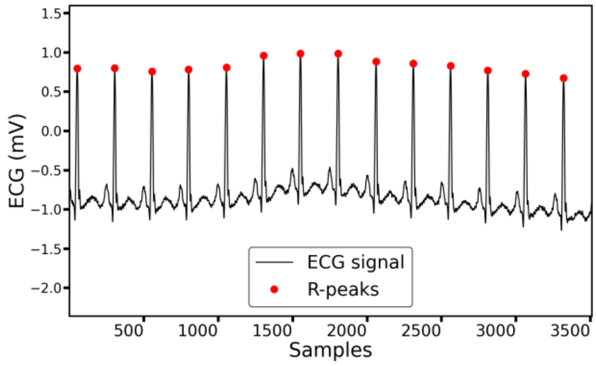


Fig. 4. ECG segments with identified R-peaks

Within each interval, the algorithm determines the maximum value of the original ECG signal. These values, along with their corresponding positions, are designated as R-peaks. Fig. 4 illustrates ECG segments with the identified R-peaks.

Remark: As observed in Fig. 4, the algorithm remains effective despite baseline drift.

2.3. Termination condition for the smoothing procedure

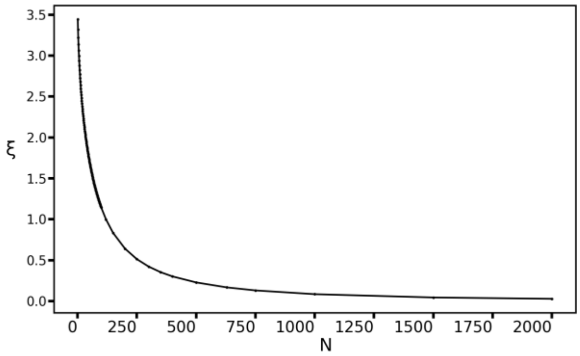
The smoothing of a function constitutes a fundamental step in data processing, particularly in the presence of noisy signals. A central problem lies in the determination of an appropriate stopping criterion for the iterative procedure. Specifically, one may inquire whether the execution of 150 iterations in the preceding subsection was sufficient. If the number of iterations is too small, the resulting function may retain excessive noise; conversely, if the number is too large, the procedure may oversmooth the signal, thereby erasing significant structural information. Let us consider the following criterion for the smoothing process at each iteration:

$$\zeta = \sum_{i=1}^{n} (H_2(\tau_{i+1}) - H_2(\tau_i))^2,$$

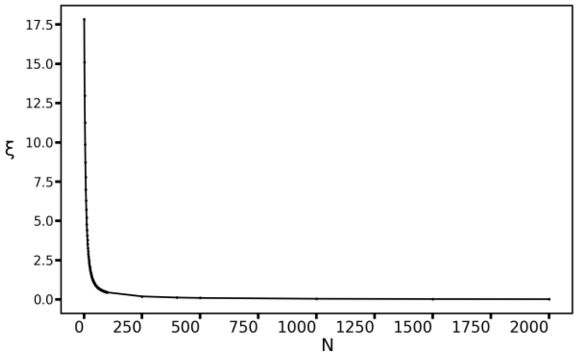
where n – denotes the number of points in the original signal. The analysis of the dependence of ζ on the number of iterations reveals a typical behaviour: initially, the sum of squared differences decreases rapidly, after which the process slows down (Fig. 5).

To determine the appropriate stopping point, several approaches can be employed. One such approach is to terminate the procedure once the difference between consecutive values of ζ becomes smaller than a prescribed threshold. Another option is to impose a maximum number of iterations, motivated by practical considerations.

ISSN 1814-4225 (print)
ISSN 2663-2012 (online)



(a)



(b)

Fig. 5. Plot of the dependence of ζ on the number of iterations performed: (a) Low noise in the signal; (b) High noise in the signal

Examination of the resulting plots indicates that the smoothing process already decelerates considerably after approximately 100 iterations.

Let us consider a fragment of an ECG signal that contains substantial noise (Fig. 6a). Figure 6 illustrates the same fragment after smoothing with different numbers of iterations. It can be observed that 100 iterations are sufficient to render the ECG signal suitable for further processing, whereas at 1000 iterations, although the characteristic peaks remain distinguishable, the signal becomes excessively smoothed.

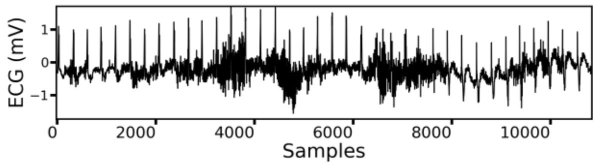
Thus, when the signal is smoothed using the method described above, empirical evidence indicates that 100 iterations are sufficient for subsequent processing, even in the presence of intense noise.

2.4. Accuracy evaluation metrics of the proposed approach

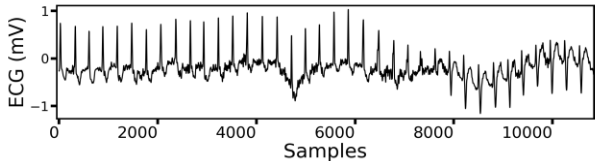
To compare the performance of the developed software with other R-peak detection algorithms, the metrics Recall (Se) and Positive Predictive Value (PPV) were calculated for each dataset. These metrics are defined by the following formulas and are adopted in the AAMI/ANSI EC38:2007 standard:

$$\mathrm{Se} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \times 100.$$

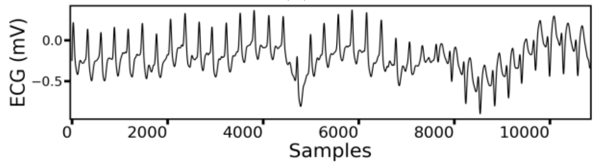
$$\mathrm{PPV} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \times 100.$$



(a)



(b)



(c)

Fig. 6. Changes in signal smoothness depending on the number of smoothing iterations: (a) without smoothing; (b) 100 smoothing iterations; (c) 1000 smoothing iterations

For these metrics, TP (True Positive) represents correctly detected R-peaks, FN (False Negative) indicates missed R-peaks, and FP (False Positive) refers to incorrectly detected R-peaks.

Additionally, the F_1 score was selected as a metric, which is the harmonic mean of Precision and Recall, used to evaluate the effectiveness of feature recognition methods. It is widely applied in assessing algorithms for detecting QRS complexes and R-peaks [23]:

$$\mathrm{F}_1\ \mathrm{score} = \frac{\mathrm{PPV} \times \mathrm{Se}}{\mathrm{PPV} + \mathrm{Se}} \times 2.$$

3. Results evaluation

3.1. Used materials

The research employed the previously mentioned MIT-BIH Arrhythmia Database, hosted on the PhysioNet platform [15]. This publicly available database is widely used in medical studies focused on analyzing ECGs to detect cardiac rhythm irregularities. It includes 48 half-hour, two-channel ECG recordings from 47 patients with diverse profiles, encompassing both inpatients under medical care and outpatients receiving ambulatory treatment during the recording period. The recordings were digitized at a 360 Hz sampling rate with an 11-bit resolution within a 10 mV range. Each signal was annotated by at least two independent cardiologists, with labels subsequently reconciled.

3.2. Experiment results

The approach proposed in this study is evaluated using files from the database [15]. As the records in the

Table 1

Comparison of the obtained results with known ones

№	Results of [25]		Calculated using the presented algor			Calculated on [26		Calculated or
	PPV(%)	Se (%)	PPV (%)	Se (%)	F_1 score (%)	Se (%)	F_1 score(%)	F_1 score(%)
100	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
101	99.84	99.84	99.89	99.89	99.89	99.84	99.92	99.84
102	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
103	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
104	99.60	99.96	97.97	100.0	98.97	99.73	99.62	99.77
105	99.42	98.73	98.61	99.73	99.17	98.95	98.98	99.07
106	99.95	99.65	100.0	100.0	100.0	99.61	99.78	99.80
107	99.95	99.90	100.0	99.9	99.97	99.67	99.67	99.92
108	99.77	99.48	97.02	99.88	98.43	99.49	99.21	99.63
109	100.0	99.92	100.0	100.0	100.0	100.0	100.0	99.96
111	100.0	99.95	100.0	99.95	99.9	99.95	99.97	99.97
112	100.0	100.	100.0	100.0	100.0	100.0	100.0	100.0
113	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
114	99.79	99.84	100.0	99.95	99.97	100.0	99,94	99.81
115	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
116	99.87	98.96	100.0	99.33	99.66	99.17	99.5	99.41
117	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
118	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
119	99.95	100.0	100.0	100.0	100.0	100.0	99.97	99.97
121	100.0	99.89	100.0	100.0	100.0	99.95	99.97	99.94
122	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
123	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
124	100.0	100.0	100.0	100.0	100.0	99.94	99.97	100.0
200	99.92	99.92	100.0	100.0	100.0	99.88	99.88	99.92
201	100.0	97.40	100.0	97.84	98.91	99.24	98.96	98.68
202	99.33	99.3	99.44	100.0	99.71	99.39	99.58	99.71
203	98.88	98.09	98.86	99.59	99.23	96.15	97.61	98.48
205	100.0	99.70	100.0	99.96	99.98	99.32	99.62	99.84
207	93.02	99.57	93.04	99.57	96.39	99.09	98.56	96.2
208	99.90	99.36	99.98	100.0	99.66	99.49	99.69	99.62
209	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
210	99.92	98.91	99.85	99.78	99.81	98.53	99.22	99.41
212	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
213	100.0	99.91	100.0	100.0	100.0	100.0	100.0	99.95
214	100.0	100.0	100.0	100.0	100.0	99.73	99.85	100.0
215	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
217	99.95	99.82	100.0	99.91	99.96	99.95	99.95	99.87
219	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
220	100.0	100.0	100.0	100.	100.0	100.0	100.0	100.0
221	100.0	99.92	99.87	100.0	99.93	99.84	99.92	99.95
222	99.92	100.0	100.0	100.0	100.0	99.8	99.88	99.96
223	100.0	99.92	100.0	100.0	100.0	99.81	99.9	99.96
228	99.21	99.71	99.41	99.56	99.48	99.85	99.7	99.48
230	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
231	100.0	99.94	100.0	100.0	100.0	100.0	100.0	99.97
232	100.0	100.0	100.0	100.0	100.0	100.0	99.72	100.0
233	100.0	99.90	99.97	99.97	99.97	99.94	99.97	99.95
234	100.0	99.96	100.0	99.96	99.98	100.0	100.0	99.98
Average	99.92	99.73	99.66	99.89	99.77	99.72	99.76	99.75

ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

mentioned database have R-peaks classified by real cardiologists, a comprehensive set of annotations is available for testing the developed algorithms. Various studies dedicated to improving peak detection algorithms report different total counts of annotated R-peaks [24 - 26] identified in the corresponding ECG database files. This discrepancy arises because different authors may select distinct subsets of peak types.

For comparison with the R-peak detection results of other authors, the results from study [25] and [26] were selected. Based on the literature review, the results presented in that publication are the most accurate within the class of threshold-based methods. Consequently, during peak detection, the same classes as those in study [25] were considered, specifically 14 peak classes, with a total count of 109,494. To compare the accuracy of peak detection, the F_1 score was chosen as the evaluation metric. The results of R-peak detection using the algorithm proposed in this study are presented in Table 1. For each file in the database, the F_1 score values were calculated for both the results obtained in the present study and those from paper [25].

In addition to the comparison with [25], the proposed algorithm was also evaluated against the results reported in study [26]. The mean F_1 score and sensitivity were computed across all database files for all compared methods. The averaged results demonstrate that the proposed algorithm achieves a higher mean F_1 score than both [25] and [26], indicating superior overall detection accuracy. Furthermore, a comparison of Sensitivity was conducted, in which the proposed algorithm likewise outperforms both referenced studies. A higher Recall value indicates that the algorithm correctly identifies a greater proportion of true R-peaks, meaning it produces fewer false negatives.

From Table 1, it can be concluded that the algorithm performs at a high level and, in almost all cases, has an improved F_1 score value. Additionally, in each of the processed files, the sensitivity value is higher than that of the algorithm from the results in [25,26], with which the comparison was made. The proposed method enabled the detection of the detection of nearly all peaks present in the cardiogram, with values falling below 99%-100% in only two cases.

4. Discussion

In most modern studies on R-peak detection in ECG, the metrics PPV, Se, and F_1 score demonstrate high and comparable results across different algorithms. However, certain database records exhibit deviations, indicating room for improvement. For example, in record 207 the PPV value is only 93.02%, which points to a significant number of false positives, where the algorithm incorrectly classifies noise as an R-peak. Although the sensitivity is high (Se = 99.57%), the false detections reduce the F_1 score. This can be explained by the presence of an anomalous segment in the signal (Fig. 7), which the algorithm mistakenly identifies as a peak. In record 203, both PPV (98.88%) and Se (98.09%) are lower than in other records, indicating that true R-peaks are being missed while false detections still occur. Similarly, in record 201 the sensitivity reaches 97.40%, whereas PPV is 100%, meaning that the algorithm failed to detect some peaks, but all detected ones were correct.

After determining the positions of the R-peaks, the algorithm computes RR-intervals - the distances between consecutive peaks. Typically, incorrectly detected peaks occur in so-called “anomalous segments” of the signal. Intervals that significantly deviate from the mean are labelled as anomalous and flagged for further verification.

For instance, in record 207 anomalous segments produce irregular RR-intervals, allowing the algorithm to identify them as “suspicious.”

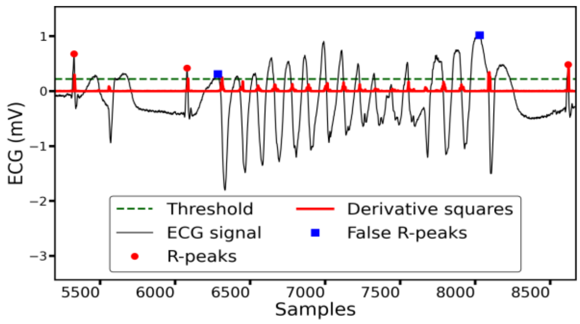


Fig. 7. An ECG fragment containing an episode representing an irregular segment of the signal (file 207 from database [15]) and the squares of the derivative

5. Conclusions

The approach proposed in this study, along with its algorithmic implementation, demonstrates high efficiency in processing ECG signals with various types of arrhythmias under different noise conditions. Testing of the algorithm on the open MIT-BIH database has shown that the accuracy of R-peak detection reaches 99%. An important feature of the algorithm is its adaptability to diverse ECG recording conditions–it operates effectively with both clean and noisy signals.

In most studies, authors do not consider certain waveform components, which are clinically significant for cardiologists, as true peaks. From a research perspective, this exclusion is justified, since these waves indeed differ from typical systolic peaks. However, it should be emphasized that such signal segments have their own medical classification and represent an important component to which physicians must pay attention. Since the intervals between these waves deviate from the norm, the algorithm correctly classifies them as “suspicious.”

Future research will focus on investigating the proposed approach on real-world ECG recordings from patients, with physicians included in the decision-making process. Studies will also be conducted on the precise localization of other informative peaks on ECG.

Contributions of authors: Mykola Yefremov – implementation of the algorithm, data preparation, conducting computational experiments, discussion of results, writing and editing; **Andrey Liashko** – data analysis and interpretation, discussion of results, conducting computational experiments; **Iurii Krak** – research concept and methodology, analysis of obtained solutions; **Oleg Stelia –** definition of research objectives, development of the proposed approach model, and validation of the obtained results.

Conflict of Interest

The authors declare no conflict of interest–financial, personal, authorship-related, or any other–that could have influenced the research process, or the results presented in this article.

Financing

This study was conducted without financial support.

Data Availability

The work has associated data in the open data repository.

Use of Artificial Intelligence

During the preparation of this work, the authors used GPT-4o and Grammarly in order to: Grammar and spelling check. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the publication’s content.

References

1. Pan, J., & Tompkins, W., J. A Real-Time QRS Detection Algorithm. *IEEE Transactions on Biomedical Engineering*, 1985, vol. BME-32, no. 3, pp. 230-236. DOI: 10.1109/TBME.1985.325532.
2. Imtiaz, M. N., & Khan, N. Pan-Tompkins++: A Robust Approach to Detect R-peaks in ECG Signals. *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Las Vegas, NV, USA, IEEE, 2022, pp. 2905-2912. DOI: 10.1109/BIBM55620.2022. 9995552.
3. Dogan, H., & Dogan, R.O. A Comprehensive Review of Computer-based Techniques for R-Peaks/QRS Complex Detection in ECG Signal, *Arch Computat Methods Eng*, 2023, vol. 30, pp. 3703–3721. DOI: 10.1007/s11831-023-09916-x.
4. Yadav, A., & Grover, N. A Review of R Peak Detection Techniques of Electrocardiogram (ECG), *Journal of engineering and technology*, 2017, vol. 8, iss. 2 pp. 115-134. Available at^ https://jet.utem.edu.my/jet/article/ view/1946. (accessed 10.11.2025).
5. Guo, T., Zhang, T., Lim, E., López-Benítez, M., Ma, F., & Yu, L. A Review of Wavelet Analysis and Its Applications: Challenges and Opportunities, *IEEE Access*, 2022, vol. 10, pp. 58869-58903, DOI: 10.1109/ACCESS.2022.3179517.
6. Krak, Iu., Stelia, O., Pashko, A., Horozov, O., & Efremov, M. Electrocardiogram Classification Using Wavelet Transformations, *IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET)*, IEEE, 2020, pp. 930-933. DOI: 10.1109/TCSET49122.2020. 235573.
7. Farrokhi, S., Dargie, W., & Poellabauer, C. Reliable peak detection and feature extraction for wireless electrocardiograms, *Computers in Biology and Medicine*, 2025, vol. 185, 9 p, DOI: 10.1016/j.compbiomed.2024.109478.
8. Bensegueni, S. A New Method for Electrocardiogram Features Extraction Using Slope Change Coefficients, *International Journal of Electronics and Telecommunications*, 2023, vol 39, iss. 1, pp. 33-39. DOI: 10.24425/ijet.2023.144328.
9. Abdullah, Al. Z.M., Thapa, K., & Yang, S.-H. Improving R Peak Detection in ECG Signal Using Dynamic Mode Selected Energy and Adaptive Window Sizing Algorithm with Decision Tree Algorithm, *Sensors*, 2021, vol. 21, article no. 6682, 17 p. DOI: 10.3390/s21196682.
10. Arora, N., & Mishra, B. Detection and classification of atrial and ventricular cardiovascular diseases to improve the cardiac health literacy for resource constrained regions, *Healthcare Technology Letters*, 2023, vol 10, iss. 3, pp. 35-52. DOI: 10.1049/htl2.12043.
11. Fariha, Z., Ikeura, R., Hayakawa, S., & Tsutsumi, S. Analysis of Pan-Tompkins Algorithm Performance with Noisy ECG Signals, *Journal of Physics Conference Series*, 2020, vol. 1532, 11 p. DOI: 10.1088/1742-6596/1532/1/012022.
12. Balta, D., & Akyemis, E. Arrhythmia Detection using Pan-Tompkins Algorithm and Hilbert Transform with Real-Time ECG Signals, *Academic Perspective Procedia*, vol. 4. iss, 1. pp. 307-315. DOI: 10.33793/acperpro.04.01.45.

13. Yen, H. T., Tiep, V. T., Hoang, V. P., Trinh, Q. K., Nguyen, H. D., Tuyen, N. T., & Sun, G. Radar-based contactless heart beat detection with a modified Pan–Tompkin's algorithm. *Biomedical Physics & Engineering Express*, 2024, vol. 11 iss. 1, 12 p. DOI: 10.1088/2057-1976/ad8c48.

14. Porr, B., & Howell, L. R-peak detector stress test with a new noisy ECG database reveals significant performance differences amongst popular detectors, *PLOS ONE*, 2019. 27 p, DOI: 10.1101/722397.

15. Moody, G. B., & Mark R. G. The impact of the MIT-BIH Arrhythmia Database. *IEEE Engineering in Medicine and Biology Magazine*, 2001, vol. 20, iss. 3, pp. 43-50, DOI: 10.1109/51.932724.

16. Goldberger, A., Amaral, L., Glass L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., Mietus, J. E., Moody, G. B., Peng, C. K., & Stanley. H. E. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation*, 2000, vol. 101, no. 23, article no. e215-e220. DOI: 10.1161/01. CIR.101.23. e215.

17. Ali, S. T. A., Kim, S., & Kim, Y.-J. Towards Reliable ECG Analysis: Addressing Validation Gaps in the Electrocardiographic R-Peak Detection, *Applied Sciences*, 2024, vol. 14 iss. 21, article 10078, DOI: 10.3390/app142110078.

18. Astre-Domínguez, C., Shmaliy, Y. S., Ibarra-Manzano, O., Muñoz-Minjares, J., & Morales-Mendoza, L. ECG Signal Denoising and Features Extraction Using Unbiased FIR Smoothing, *BioMed Research International*, 2019, article no. 2608547. 16 p. DOI: 10.1155/2019/2608547.

19. Shmaliy, Y., & Morales-Mendoza, L. FIR Smoothing of Discrete-Time Polynomial Signals in State Space. *IEEE Transactions on Signal Processing,* 2010, vol. 58, pp. 2544-2555. DOI: 10.1109/TSP.2010. 2041595.

20. Niedźwiecki, M., J., Ciołek, M., Gańcza, A., & Kaczmarek, P. Application of regularized Savitzky–Golay filters to identification of time-varying systems, *Automatica, 2*021, vol. 133. 9 p. DOI: 10.1016/j.automatica.2021.109865.

21. Chatterjee, S., Thakur, R. S., Yadav, R. N., Gupta, L., & Raghuvanshi, D., K. Review of noise removal techniques in ECG signals, *IET Signal Processing*, 2020, vol. 14, iss, 9, pp 569-590, DOI: 10.1049/iet-spr.2020.0104f.

22. Ardeti, V. A., Kolluru, V. R., Varghere, G. T., & Patjoshi, R. K. An overview on state-of-the-art electrocardiogram signal processing methods: Traditional to AI-based approaches, *Expert Systems with Applications*, 2023, vol. 217. DOI: 10.1016/j.eswa.2023.119561.

23. Krak, I., Pashko A., Stelia O., Barmak, O., & Pavlov, S. Selection Parameters in the ECG Signals for Analysis of QRS Complexes. *1st International Workshop on Intelligent Information Technologies and Systems of Information Security*, *InteIITSIS*, 2020, vol. 2623. 13 p. Available at https://ceur-ws.org/Vol-2623/paper1.pdf. (accessed 15.12.2025).

24. Association for the Advancement of Medical Instrumentation (R2020) ANSI/AAMI EC57:2012; *Testing and Reporting Performance Results of Cardiac Rhythm and ST Segment Measurement Algorithms.* Available at: https: //webstore.ansi.org/standards/aami/ ansiaamiec572012r2020 (accessed 11.09.2025).

25. Zhai, D., Bao, X., Xi Long, X., Ru, T., & Zhou, G. Precise detection and localization of R-peaks from ECG signals, *Mathematical Biosciences and Engineering*, 2023, vol. 20, iss. 11, pp. 19191-19208. DOI: 10.3934/mbe.2023848.

26. Hsieh, F., & Chen, T. -L. A Novel R-Peak Detection Algorithm. IEEE Access, 2025, vol. 13, pp. 210351-210359. DOI: 10.1109/ACCESS.2025.3643153.

Received 08.06.2025, Received in revised form 06.12.2026
Accepted date 15.01.2026, Published date 21.01.2026

ЕФЕКТИВНИЙ ПОРОГОВИЙ МЕТОД ВИЗНАЧЕННЯ R-ЗУБЦІВ В ЕЛЕКТРОКАРДІОГРАМІ

М. С. Єфремов, А. В. Ляшко, Ю. В. Крак, О. Б. Стеля

Автоматизація виявлення R-зубців в електрокардіограмі (ЕКГ) є одним з ключових етапів аналізу серцевого ритму, що використовується для аналізу частоти серцевих скорочень, оцінки їх варіабельності, діагностики аритмій тощо. На складність знаходження зубців впливає якість запису ЕКГ, особливо наявність шуму та артефактів. У дослідженні запропоновано метод та здійснено його програмну реалізацію для виявлення R-зубців, що ґрунтується на згладжуванні сигналу із подальшим його диференціюванням та використання порогового підходу. Особливістю методу є забезпечення неперервності функції та її похідної, що дозволяє з великою точністю виявити зубці на складних ділянках сигналу. Метод орієнтований на застосування в умовах обмежених обчислювальних ресурсів для портативних і вбудованих систем моніторингу. **Метою** дослідження є розроблення обчислювально ефективного та точного методу автоматичного виявлення R-зубців в ЕКГ-сигналі, орієнтованого на персоніфікований підхід до пацієнта.

Основна увага приділяється стійкості до шуму та морфології QRS-комплексу, а також здатності алгоритма працювати в умовах обмежених ресурсів та у режимі реального часу. Запропонований метод базується на поетапній обробці ЕКГ-сигналу. Для впевненості у доцільності підходу та для порівняння ефективності розробленого методу було проаналізовано відомі дослідження у даній сфері. **Тематика** дослідження полягає у обробленні сигналів з ЕКГ та аналізі проблем виявлення R-зубців на ЕКГ знятих за допомогою Холтерів шляхом побудови згладженого неперервного сигналу на основі дискретних даних, отриманих у процесі оцифрування. Дискретність початкового сигналу ускладнює диференціювання та точне визначення характерних точок, зокрема R-зубців, що відіграють ключову роль у діагностиці серцевих захворювань. **Метод** кусково-поліноміального згладжування вхідних даних з метою отримання неперервної функції спеціального вигляду дозволив не лише отримати неперервну функцію, а й забезпечити неперервність її першої похідної, що критично важливо для подальшого аналізу поведінки сигналу та визначення точок з різкою зміною кривизни сигналу. **Наукова новизна** цього дослідження полягає у використанні кусково-поліноміального наближення другого порядку для представлення сигналу ЕКГ. Цей підхід дозволяє зменшити шум в сигналі та представити дискретний сигнал у вигляді функції неперервної разом зі своєю першою похідною та коректно обчислювати її похідну. **Результатом** є виявлення R-зубців шляхом аналізу похідної згладженого сигналу: було виділено ділянки з різкими змінами, які характерні для QRS-комплексу та побудовано ітеративну схему згладжування, кількість ітерацій у якій визначено запропонованим критерієм зупинки. Запропонований метод було реалізовано програмно та протестовано на даних, взятих з відкритої бази даних MIT-BIH Arrhythmia Database що містить понад 60 записів та більше ста тисяч анотованих R-зубців. Результати порівняно з дослідженнями інших авторів, використовуючи метрику F_1 score на базі стандартних метрик точності та чутливості. **Висновки:** в дослідженні пропонується ефективний та адаптивний підхід до обробки ЕКГ-сигналів, що забезпечує надійне виявлення R-зубців за умов значного шуму, дрейфу базової лінії та індивідуальних відмінностей пацієнтів. Отримані результати продемонстрували високі значення за метрикамb: до 99.5% (середнє понад 99.1%), F_1 score: стабільно перевищує 99%, у деяких записах – 100%. Таким чином запропонований підхід є конкурентоспроможним, демонструє високу точність виявлення R-зубців при аритміях, інверсіях зубців, нестандартній морфології QRS-комплексу тощо. Результати дослідження можуть бути використані у практиці аналізу ЕКГ, зокрема у побудові автоматизованих систем діагностики або попередньої обробки сигналу перед застосуванням класифікаційних методів.

Ключові слова: електрокардіограма; кусково-поліноміальне наближення; R-зубці; згладжування; диференціювання ЕКГ; MIT-BIH.

Єфремов Микола Сергійович – асист. каф. теоретичної кібернетики, Київський національний університет імені Тараса Шевченка, Київ, Україна

Ляшко Андрій Володимирович – асист. каф. теоретичної кібернетики, Київський національний університет імені Тараса Шевченка, Київ, Україна

Крак Юрій Васильович – чл.-кор. НАН України, д-р фіз-мат. наук, проф., зав. каф. теоретичної кібернетики, Київський національний університет імені Тараса Шевченка; провідний науковий співробітник, Інститут кібернетики імені В. М. Глушкова НАН України, Київ, Україна.

Стеля Олег Борисович – канд. фіз.-мат. наук, доц., Київський національний університет імені Тараса Шевченка; старш. наук. співроб. Інституту кібернетики імені В. М. Глушкова НАН України, Київ, Україна.

Mykola Yefremov – Assistant of the Theoretical Cybernetics Department, Taras Shevchenko National University of Kyiv, Kyiv, Ukraine,
e-mail: yefremov@knu.ua, ORCID: 0000-0001-8698-3957, Scopus Author ID: 57217113343.

Andrey Liashko – Assistant of the Theoretical Cybernetics Department, Taras Shevchenko National University of Kyiv, Kyiv, Ukraine,
e-mail: andrey.liashko@knu.ua, ORCID: 0009-0006-1335-2314.

Iurii Krak – Corresponding Member NAS of Ukraine, Doctor of Physical and Mathematical Sciences, Professor, Head of the Theoretical Cybernetics Department, Taras Shevchenko National University of Kyiv, Leading researcher of Glushkov Cybernetics Institute NAS of Ukraine, Kyiv, Ukraine,
e-mail: iurii.krak@knu.ua, ORCID: 0000-0002-8043-0785, Scopus Author ID: 6602577533.

Oleg Stelia – Candidate of Physical and Mathematical Sciences, Associate Professor, Taras Shevchenko National University of Kyiv, Senior Researcher Fellow of Glushkov Cybernetics Institute NAS of Ukraine, Kyiv, Ukraine,
e-mail: oleg.stelia@knu.ua, ORCID: 0000-0002-1453-501X, Scopus Author ID: 57201773676.