

DOI: 10.17725/rensit.2024.16.053

## Классификация сигналов ЭКГ на основе цифровой обработки сигналов, технического выбора функций и радиочастотного классификатора

<sup>1</sup>Анас Фуад Ахмед, <sup>2</sup>Хадил Н. Абдуллах, <sup>1</sup>Бараа М. Албакер

<sup>1</sup>Иракский университет, факультет электротехники, <http://www.aliraqia.edu.iq/>  
Адхамия, Багдад 6029, Ирак

<sup>2</sup>Технологический университет, электротехнический факультет, <http://www.uotechnology.edu.iq/>  
Багдад 19006, Ирак

*E-mail:* [anas.abmed@aliraqia.edu.iq](mailto:anas.abmed@aliraqia.edu.iq), [30002@uotechnology.edu.iq](mailto:30002@uotechnology.edu.iq), [baraamalbakker@aliraqia.edu.iq](mailto:baraamalbakker@aliraqia.edu.iq)

Поступила 11.09.2023, рецензирована 18.09.2023, принята 25.09.2023, опубликована 15.03.2024.

Представлена действительным членом РАЕН А.С. Дмитриевым

**Аннотация:** Определение оптимальной интеграции между функциями и классификаторами оказывает существенное влияние на производительность автоматических систем диагностики сердечбиения. Актуальность этого проявляется при работе с критически важными приложениями, которые содержат устройства с ограниченными ресурсами и требуют точных и быстрых классификаторов сердечбиения, чтобы помочь врачу поставить точный и быстрый диагноз заболеваний сердца. С целью решения этой задачи в данной статье представлен новый подход к выбору оптимальных характеристик сигнала ЭКГ путем использования радиочастотного классификатора (RF), следуя методу разделения сигналов ЭКГ между пациентами и следуя инструкциям Ассоциации по развитию медицинского инструментария (The Association for the Advancement of Medical Instrumentation, AAMI). Характеристики были выбраны на основе концепции "взаимного информационного ранжирования (Mutual Information Ranking, MIR)". Представленная структура является комплексной с точки зрения эффективного выполнения всех необходимых процессов, начиная от обработки цифрового сигнала ЭКГ, сегментации, извлечения признаков, выбора признаков и заканчивая классификацией ЭКГ. Результаты экспериментов показывают, что признаки, соответствующие нормализованной ширине QRS и нормализованным интервалам RR, являются наиболее влиятельными признаками в классификации сердечного ритма. Все тесты проводились с использованием реальных сигналов ЭКГ, взятых из базы данных "MIT-BIH" (MIT-BIH-ARR-DB) по аритмиям. Предложенная схема достигла следующих показателей F1: 91.02%, 73.17% и 98.04% по классификации желудочковых эктопических комплексов (V или VEB), наджелудочковых эктопических комплексов (S или SVEB) и нормальных комплексов (N или NB), соответственно. Общая точность составила 96.26%. Несмотря на свою относительную простоту и использование небольшого количества функций, предлагаемый подход превосходит большинство известных современных решений.

**Ключевые слова:** взаимная информация; выбор функции; аритмия сердца; дискретное вейвлет-преобразование; классификация ЭКГ; радиочастотный классификатор

УДК 53.075.8:57.043

*Для цитирования:* Анас Фуад Ахмед, Хадил Н. Абдуллах, Бараа М. Албакер. Классификация сигналов ЭКГ на основе цифровой обработки сигналов, технического выбора функций и радиочастотного классификатора. РЭНСИТ: Радиозлектроника. Наносистемы. Информационные технологии, 2024, 16(1):53-66. DOI: 10.17725/rensit.2024.16.053.

# A Robust and Efficient Scheme for ECG Signal Classification Based on Digital Signal Processing, Feature Engineering Selection, and Random Forest Classifier

Anas Fouad Ahmed, Baraa M. Albaker

Al-Iraqia University, College of Engineering, Electrical Engineering Department, <http://www.aliraqia.edu.iq/>  
Al Adhmia-Haiba Khaton, Baghdad 6029, Iraq

E-mail: [anas.abmed@aliraqia.edu.iq](mailto:anas.abmed@aliraqia.edu.iq), [baraamalbaker@aliraqia.edu.iq](mailto:baraamalbaker@aliraqia.edu.iq)

Hadeel N. Abdullah

University of Technology, Electrical Engineering Department, <http://www.uotechnology.edu.iq/>  
Al wehada-Neighborhood, Baghdad 19006, Iraq

E-mail: [30002@uotechnology.edu.iq](mailto:30002@uotechnology.edu.iq)

Received September 11, 2023, peer-reviewed September 18, 2023, accepted September 25, 2023, published March 15, 2024.

**Abstract:** Determining the optimal integration between features and classifiers has a significant effect on the performance of automatic heartbeat diagnostic systems. This importance stands out when dealing with critical applications that contain limited resources devices and require accurate and fast heartbeat classifiers to help the doctor make an accurate and quick diagnosis of heart diseases. Aiming at this task, this paper introduces a novel approach for choosing the optimal features of the ECG signal to be used with the Random Forest (RF) classifier following the inter-patient method for ECG signals division and obeying the instructions of the "Association for the Advancement of Medical Instrumentation (AAMI)." The features were chosen based on the concept of "Mutual Information Ranking (MIR)." The presented framework is comprehensive in terms of performing all the necessary processes efficiently, starting from ECG digital signal processing, segmentation, feature extraction, feature selection, and ending with ECG classification. The results of the experiments demonstrate that features corresponding to the normalized QRS width and the normalized RR intervals are the most influential features in the heartbeat classification. All tests were conducted using real ECG signals taken from the "MIT-BIH" Arrhythmia Database (MIT-BIH-ARR-DB). The suggested scheme attained the following F1-scores: 91.02%, 73.17%, and 98.04% in the classification of the Ventricular Ectopic Beats (V or VEB), Supraventricular Ectopic Beats (S or SVEB), and Normal Beats (N or NB), respectively. The overall accuracy was 96.26%. Despite its relative simplicity and reliance on few features, the proposed approach outperforms most of the reported state-of-the-art.

**Keywords:** Mutual Information; Feature Selection; Cardiac Arrhythmia; Discrete Wavelet Transform (DWT); ECG classification; Random Forest (RF)

UDC 53.075.8:57.043

*For citation:* Anas Fouad Ahmed, Hadeel N. Abdullah, Baraa M. Albaker. A Robust and Efficient Scheme for ECG Signal Classification Based on Digital Signal Processing, Feature Engineering Selection, and Random Forest Classifier. *RENSIT: Radioelectronics. Nanosystems. Information Technologies*, 2024, 16(1):53-66e. DOI: 10.17725/j.rensit.2024.16.053.

## СОДЕРЖАНИЕ

### 1. ВВЕДЕНИЕ (55)

### 2. ПРЕДЛАГАЕМАЯ СТРУКТУРА (56)

#### 2.1. СБОР СИГНАЛОВ ЭКГ (56)

#### 2.2. ПРЕДВАРИТЕЛЬНАЯ ОБРАБОТКА ЦИФРОВОГО СИГНАЛА (57)

#### 2.3. ОБНАРУЖЕНИЕ ПИКА R (58)

#### 2.4. ОТБОР СЕРДЦЕБИЕНИЙ (59)

### 2.5. ФУНКЦИИ ОПИСАНИЯ КОМПЛЕКСА QRS (59)

### 2.6. ВЫБОР ФУНКЦИЙ (60)

### 2.7. КЛАССИФИКАЦИЯ СЕРДЦЕБИЕНИЙ (60)

### 2.8. ПОКАЗАТЕЛИ ОЦЕНКИ КЛАССИФИКАЦИИ (61)

### 3. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ (61)

### 4. ЗАКЛЮЧЕНИЕ (63)

### ЛИТЕРАТУРА (64)

## 1. ВВЕДЕНИЕ

Сигнал электрокардиограммы (ЭКГ) состоит из набора периодических сердечных сокращений, на базе которых формируются различные схемы классификации сердечных ритмов. Согласно информации ААМІ (The Association for the Advancement of Medical Instrumentation – Ассоциация развития медицинского оборудования) существующие подходы классифицируют сердечные ритмы, различая следующие их категории: слитные ритмы (F или FB), неклассифицируемые или стимулированные сокращения (Q или QB), VEB, SVEB и NB. Необходимо обнаружить и подсчитать количество ненормальных эктопических сокращений, чтобы избежать опасных заболеваний сердца. Например, повторение SVEB связано с опасным заболеванием "фибрилляция предсердий", также рецидив VEB является исключительно индикаторным, поскольку его можно использовать как предиктор опаснейшего заболевания "сердечная недостаточность", которое часто приводит к смерти [1-15].

В последние годы исследователи приложили немало усилий для решения проблемы автоматической диагностики или классификации сердечных сокращений, предложив различные методы выполнения этой задачи. Как правило, эти методы можно разделить на два подхода: первый основан на стратегии глубокого обучения, а другой – на разработке функций. Некоторые исследователи применили глубокие нейронные сети (DNN) для классификации сигналов ЭКГ [16-18], но DNN испытывает избыточность параметров и сложность вычислений. Здесь важно отметить, что аналитическая классификация, основанная на разработке признаков, может помочь кардиологам в точной диагностике и точном составлении подходящих графиков терапии, используя преимущества понятных извлеченных признаков. С другой стороны, алгоритмы глубокого обучения не могут предоставить кардиологам интерпретируемые функции и не могут анализировать влияние выбора функций на эффективность классификации ЭКГ, поскольку эти функции неявно фиксируются глубокими уровнями сети. Одно из наиболее важных исследований схем классификации

сердечных ритмов было представлено Luz et al. [19]. В настоящее время большой интерес вызывают методы машинного обучения (ML), и, пожалуй, самым известным из них является метод RF, в котором набор классификаторов создается в виде графа - "ансамбля" деревьев решений (DT) [20]. Вход классифицируется по ансамблю (лесу) по мажоритарному принципу [21]. Поэтому RF-классификатор считается чрезвычайно эффективным и рекомендован для устройств с ограниченными ресурсами, предназначенных для работы в режиме реального времени [22-25]. Emanit [26] применил DWT (дискретное вейвлет-преобразование) к сигналам ЭКГ и использовал полученные коэффициенты в качестве признаков для обучения радиочастотного классификатора. Ламедо и др. [27] представлен простой классификатор ЭКГ, который классифицирует SVEB и VEB на основе коэффициентов интервалов DWT и RR. Аликович и Субаси [28] предложили использовать распределение коэффициентов DWT для извлечения различных характеристик, которые вводятся в RF для целей классификации сердечных сокращений. Гутьеррес и др. [29] классифицировали сигналы ЭКГ с использованием квадратичного вейвлет-преобразования (QWT) и нейронных сетей. Ганеш и Кумарасвами [30] предложили метод классификации сигналов ЭКГ на основе интервалов RR и признаков, извлеченных с помощью дискретного косинусного преобразования (DCT) сигналов ЭКГ для обучения радиочастотного классификатора. Парк и др. [31] использовали коллекции различных функций во временной области с радиочастотным классификатором для классификации пульса. Аравинд и др. [32] предложили алгоритм классификации пульса с использованием сверточной нейронной сети (CNN) и непрерывного вейвлет-преобразования (CWT). Юаньлу и др. [3] классифицировали сигнал ЭКГ на основе глубокого остаточного CNN. Вышеупомянутые исследования достигли превосходной точности (ACC); однако существуют некоторые ограничения, например, некоторые из них: соблюдение рекомендаций ААМІ относительно типа сердечного ритма, которое необходимо обнаружить и классифицировать, предложена надежная модель, которая может быть применена непосредственно к

несбалансированным данным или сигналам и использованы другие показатели для оценки качества классификации, такой как показатель F1, помимо положительной прогнозируемости (PPr) и чувствительности (Sen), которые могут ухудшиться из-за доминирования класса большинства. Самое главное, что ранее (насколько известно авторам) не было исследований, посвященных использованию радиочастотного классификатора, подчиняющегося межпациентному режиму. В этом режиме сигналы разделяются на две группы: первая используется на этапе обучения, а вторая используется на этапе тестирования, при этом между двумя группами нет пересечения (нет сигналов, полученных от одного и того же пациента в обе группы), что приводит к более достоверной оценке работы классификатора. В этом исследовании мы использовали эффективный радиочастотный классификатор для классификации сигналов ЭКГ в соответствии с рекомендациями ААМІ и межпациентным режимом. Ниже приведены результаты этой работы:

- Мы представляем комплексную, надежную и эффективную основу для классификации сигналов ЭКГ на основе улучшенного радиочастотного классификатора, принимая во внимание всю необходимую цифровую обработку и анализ сигналов для собранных сигналов и придерживаясь инструкций ААМІ и межпациентного режима.
- Предложен простой и эффективный механизм обнаружения R-пикулов на основе тригонометрии.
- Представлены новые соображения по оценке нормализованных признаков.
- Оптимизация производительности RF-классификатора без необходимости

сложных вычислений с помощью MIR для получения оптимальной и сокращенной группы функций из большого набора функций, содержащей наиболее важные функции, предложенные в предыдущих исследованиях.

- Успешное решение проблемы несбалансированных данных сигналов ЭКГ, когда аномальных сигналов намного меньше, чем нормальных сигналов. Это достигается за счет использования RF-классификатора, который фактически состоит из ансамбля нескольких DT, чувствительных к несбалансированным классам, и принимает унифицированную случайную выборку (URS) из данных сигнала ЭКГ со стратегией замены.
- Понятный анализ сигнала ЭКГ достигается за счет предоставления наборов функций из различных сигналов ЭКГ, которые необходимы для медицинских целей.

## 2. ПРЕДЛАГАЕМАЯ СТРУКТУРА

Полная блок-схема предлагаемой структуры показана на **Рис. 1**. Она состоит из следующих этапов: сбор сигналов ЭКГ, предварительная обработка цифровых сигналов, учет сердцебиений, извлечение признаков, выбор признаков, классификация сердцебиений и оценка классификатора. Каждый из этих этапов более подробно описан в последующих разделах.

### 2.1. СБОР СИГНАЛОВ ЭКГ

Сигналы ЭКГ были собраны из базы данных MIT-BIH-ARR-DB по аритмиям, которая доступна в [34]. Она хорошо известна и используется многими исследователями, что облегчает процесс сравнения предложенной схемы с предыдущими работами. В **Таблице 1** кратко описаны эти сигналы.

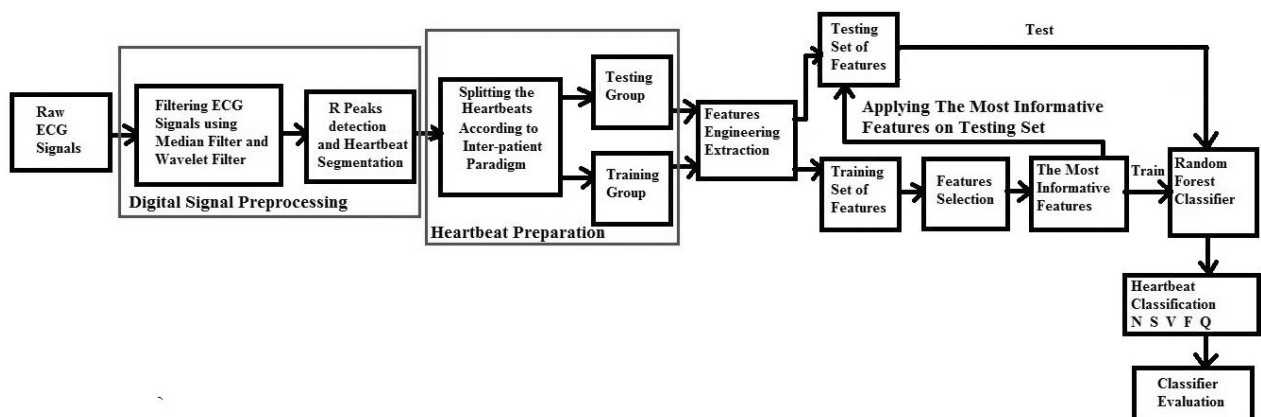


Рис. 1. Блок-схема предлагаемой структуры.

**Таблица 1**  
Описание собранных сигналов ЭКГ

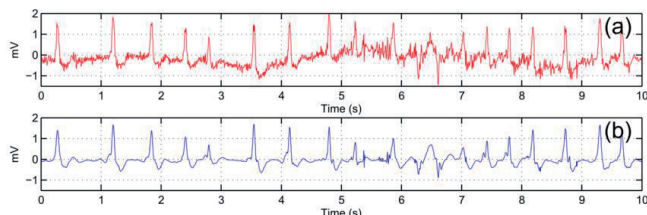
Параметр	Значение
Число записей	48
Продолжительность (мин.)	30
Частота дискретизации (Гц)	360
Проводники	Одиночный (MLII)
Число сердцебиений	109,494
Число типов сердцебиений	15
Пол участников (%)	53% мужчины, 47% женщины
Возрастной диапазон (лет)	23-89

Процесс перемаркировки собранных сигналов проводился для получения пяти классов сердцебиений в соответствии с инструкциями AAMI, как показано в **Таблице 2**.

**2.2. ПРЕДВАРИТЕЛЬНАЯ ОБРАБОТКА ЦИФРОВОГО СИГНАЛА**

На практике сигналы ЭКГ во время их получения или передачи обычно искажаются двумя основными типами шумов: "блужданием базовой линии" (BW) и белым гауссовским шумом (WGN) [35]. Эти шумы необходимо по возможности уменьшить, чтобы избежать их негативного влияния на качество сигналов и эффективность классификации. **Рис. 2** демонстрирует пример сигнала ЭКГ, загрязненного BW и WGN.

Для получения сигналов ЭКГ, скорректированных по базовой линии, каждый сигнал проходил через два медианных фильтра (ширина первого составляет 200 мс, а ширина второго – 600 мс). Затем он вычитается из



**Рис. 2.** Пример сигнала ЭКГ, искаженного BW и WGN.

необработанного сигнала ЭКГ. После этого для устранения WGN берется дискретное вейвлет-преобразование DWT сигнала, скорректированное по базовой линии, для разложения его на различные диапазоны частот. Вейвлет-фильтр Добеши-6 (DB6), четыре уровня декомпозиции, схема мягкой пороговой обработки (определенная в уравнении (1)) и универсальная стратегия определения порога (заданная в уравнении (2)) являются оптимальными для удаления WGN из сигнала ЭКГ.

$$S_{thresholding}(De) = \text{sign}(De)(|De| - t_{hr}) \quad (1)$$

for  $|De| > t_{hr}$ , and 0 for  $|De| \leq t_{hr}$ ,

где  $De$  – коэффициенты детализации;  $t_{hr}$  – пороговое значение.

$$t_{hr} = \text{SDN} \cdot \sqrt{2 \log(\text{No.Samples})} \quad (2)$$

$$\text{SDN} = \text{Median}(D) / 0.6745$$

где SDN – стандартное отклонение шума.

DB6 характеризуется следующими особенностями [35]:

- Его форма очень близка к сигналу ЭКГ, что важно для идеальной реконструкции.

**Таблица 2**

Перемаркировка MIT-BIH-ARR-DB согласно инструкциям AAMI

Классы AAMI с пояснениями		MIT-BIH ритмы с маркировками
Нормальные ритмы (N)	N	Нормальный ритм
	L	Блокада левой ножки пучка Гиса
	R	Блокада правой ножки пучка Гиса
	e	Ускользящий предсердный ритм
	j	Узловой (соединительный) выскальзывающий комплекс
Суправентрикулярный эктопический комплекс (SVEB)	A	Предсердная экстрасистолия
	a	Аберрированная предсердная экстрасистолия
	J	Узловая (соединительная) экстрасистолия
	S	Суправентрикулярная экстрасистолия
Желудочковая эктопия (VEB)	V	Преждевременное сокращение желудочков
	E	Ускользящий желудочковый ритм
Слияние ритмов (F)	F	Слияние желудочкового и нормального ритма
Неизвестные ритмы (Q)	I	Ритмичный ритм
	f	Совмещение ритмичного и нормального ритма
	Q	Неклассифицированный ритм



Рис. 3. Форма вейвлет-фильтра DB6.

- Подходит для нестационарных сигналов из-за коротких моментов исчезновения.
- Он имеет относительно низкую стоимость вычислений из-за своей ортогональности.

На Рис.3 показана форма вейвлет-фильтра DB6.

### 2.3. ОБНАРУЖЕНИЕ ПИКА R

Это исследование предлагает простую и эффективную стратегию обнаружения пика R путем оценки углов в градусах между последовательными выборками амплитуды сигнала с последующим применением динамического порога к этим углам и временному окну, как описано в уравнениях (3), (4), (5), (6) и (7) соответственно:

$$\theta(k) = \tan^{-1}\left(\frac{V}{W}\right) = \tan^{-1}\left[\frac{\beta \times |\delta(k) - \delta(k-1)|}{360 \times \text{Sampling frequency}}\right] \quad (3)$$

$$\beta = \begin{cases} 500, & \text{initial value} \\ 1000, & \left(\frac{V}{W}\right) < 57.5 \text{ or } \theta(k) < 89^\circ \text{ during } 2s \text{ interval} \\ 500, & \left(\frac{V}{W}\right) < 117 \text{ or } \theta(k) > 89.5^\circ, \end{cases} \quad (4)$$

где  $\beta$  – коэффициент увеличения,  $\delta(k)$  – k-я выборка амплитуды.

Необходимо отметить, что  $\beta$  в уравнении (4) настраивается таким образом, чтобы обеспечить обнаружение биений низкой амплитуды. Рис. 4 иллюстрирует эту концепцию: при частоте дискретизации = 360 Гц видно, что диапазон углов сохраняется между  $80^\circ$  и  $90^\circ$ . Динамическое пороговое значение в градусах  $D(k)$  варьируется в соответствии с уравнением (5), разобраться с изменениями сердечного ритма в зависимости от различных случаев заболевания пациентов и условий регистрации. Значения  $\theta(k)$  и  $D(k)$  сравниваются, затем временное окно  $G(k)$  используется для обнаружения пика R.

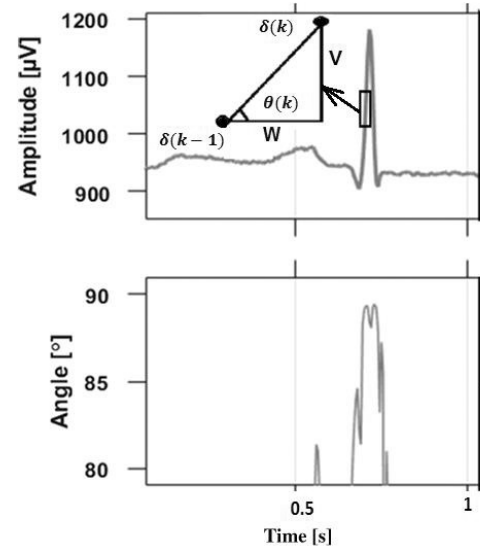


Рис. 4. Оценка углов сердцебиения на основе тригонометрии.

$$D(k+1) = \begin{cases} 0, & \text{initial value} \\ \theta(k) - Y_1, & \|I\| \theta(k) > D(k) + Y_1 \\ D(k), & \|I\| D(k) > \theta(k) \leq D(k) + Y_1, \beta t = 0 \\ D(k) - (Y_2 \times \beta t), & \theta(k) \leq D(k), \beta t = \beta t + 1 \\ 80, & |D(k)| \leq 80, \end{cases} \quad (5)$$

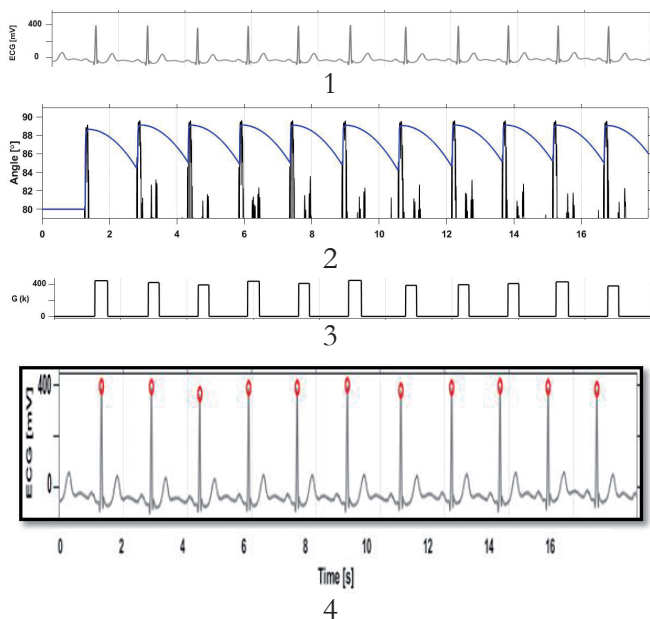
где  $Y_1$  и  $Y_2$  эмпирически установлены равными  $0.4^\circ$  и  $0.00012^\circ$ , последовательно, чтобы получить наименьшую ошибку обнаружения, как показано в Приложении А.

$$G(k) = \begin{cases} \text{Signal amplitude}, & 0 \leq \beta t \leq Y_2 \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

$$Y_2 = \begin{cases} 0.25s, & \text{initial value} \\ 0.450s \times F_s, & \text{Mean of last 8 RR intervals} \geq 0.68s \\ 0.25s \times F_s, & \text{otherwise.} \end{cases} \quad (7)$$

После оценки временного интервала  $G(k)$  проводится сравнение между |локальный максимум| и |локальный минимум| проводится амплитуда сигнала в пределах  $G(k)$ ; самый высокий из них указывается как точка пика R, как показано на Рис. 5. Сильные стороны предлагаемой стратегии можно резюмировать следующим образом. Уравнение (1) гарантирует, что связь между  $\theta(k)$  и  $|\delta(k) - \delta(k-1)|$  является нелинейной, а  $\theta(k)$  ограничен фиксированным диапазоном ( $0^\circ$  и  $90^\circ$ ). Другими словами,  $\theta(k)$  невосприимчив к различным изменениям комплекса QRS. Более того, член  $\tan^{-1}$  делает вычисление  $\theta(k)$  простым и быстрым.

После выполнения процесса фильтрации и обнаружения пика R выполняется процесс сегментации. Каждый сигнал ЭКГ сегментируется на несколько сердечных



**Рис. 5.** Наглядный пример предлагаемого подхода к обнаружению R-пиков. (1). Дешумированный сигнал, (2). Углы (черный) и динамическое определение порога (синий), (3). Временное окно  $G(k)$ , (4). Обнаружение и маркировка пиков R.

сокращений в зависимости от местоположения пиков R. Сердцебиение определяется путем указания 92 точек выборки перед пиком R и 146 точек выборки после пика R (т.е. длина каждого сердцебиения составляет 238 выборок).

**2.4. ОТБОР СЕРДЦЕБИЕНИЙ**

Сегментированные сердцебиения разделяются (согласно межпациентной схеме [36]) на две группы: обучающую группу, возникающую в результате фильтрации и сегментации сигналов набора 1, и тестирующую группу, возникающую в результате фильтрации; и сегментация сигналов набора 2. Пересечения между двумя наборами нет (они поступают от разных пациентов). Записи для набора 1 и набора 2 показаны в **Таблице 3**. Согласно стандартам AAMI [37], четыре записи, содержащие ритмические сокращения, удаляются.

**2.5. ФУНКЦИИ ОПИСАНИЯ КОМПЛЕКСА QRS**

В этой статье было рассмотрено большое количество функций, и они представляют собой наиболее примечательные восемьдесят

**Таблица 3**

Сигналы набора 1 и набора 2 с соответствующими записями

Сигналы	Записи из MIT-BIH-ARR-DB
Сигналы 1 набора	101,106,108,109,112,114,115,116,118,119,122,124,201,203,205,207,208,209,215,220,223,230
Сигналы 2 набора	100,103,105,111,113,117,121,123,200,202,210,212,213,214,219,221,222,228,231,232,233,234

пять функций, большинство из которых были рассмотрены в предыдущих работах [31,37-41], и они заключаются в следующем:

- **Коэффициенты DWT.** DWT способно извлекать подробную информацию как из частотной, так и из временной области, что делает его идеальным для представления ЭКГ. В этом исследовании материнский вейвлет Добеши первого порядка (DB1) с тремя уровнями разложения использовался для введения двадцати трех признаков.
- **Коэффициенты базисных функций Эрмита (HBF).** Коэффициенты HBF использовались в качестве характеристик, описывающих структуру сигнала ЭКГ. Точки выборки, расположенные за 250 мс до и после каждого пика R, представляют собой сегмент биения. Параметры оценки коэффициентов HBF описаны в [38]. Порядок HBF был зафиксирован равным двадцати, а переменная ширины оценивалась для уменьшения ошибки реконструкции (насколько это возможно) для каждого удара [37]. Для получения пятнадцати признаков использовались три, четыре и пять HBF.
- **Статистика высшего порядка (HOS).** Она позволяет фиксировать небольшие изменения сигналов ЭКГ [39]. Каждое сердцебиение было разбито на 5 интервалов в каждом; статистики третьего и четвертого порядка (асимметрия и эксцесс) были определены для создания десяти признаков. Диапазон переменных задержки, которые сосредоточены на пике R, составляет от -250 до 250 мс, как описано в [39].
- **Длительность комплексов QRS и их нормализация.** Они включают в себя следующие пять длительностей и их нормализаций (десять признаков): общая длительность QRS (QRS<sub>0</sub>), длительность QRS на среднем значении пика R (QRS<sub>m</sub>), длительность QRS на четверти значения пика R (QRS<sub>q</sub>), длительность между зубцами Q и S (dQS), длительность между зубцом P и начальной точкой комплекса QRS (dPQRS). Эти особенности продемонстрированы на **Рис. 6**. Нормализацию вышеупомянутых функций можно получить путем деления значения функции на среднее ее значение за последние тридцать два такта. Подробное объяснение механизма извлечения представлено в **Приложении В**.
- **Евклидовы расстояния.** Эти четыре характеристики зависят от евклидовых

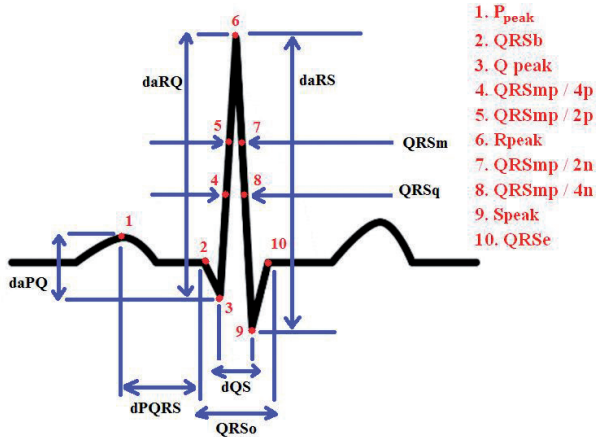


Рис. 6. Демонстрация особенностей длительности и различий амплитуд комплекса QRS.

расстояний между пиком R и четырьмя точками биения, которые представляют значения амплитуды в нескольких выборках (как предложено в [37]): максимальная амплитуда (удар [0, 40]), минимальная амплитуда (удар [75, 85]), минимальная амплитуда (удар [95, 105]) и максимальная амплитуда (удар [150, 180]).

- **Разница амплитуд и их нормализация.** Они состоят из следующих семи амплитуд и их нормировок (четырнадцати признаков): разница в амплитуде между зубцами R и Q ( $daRQ$ ), разница в амплитуде между зубцами P и Q ( $daPQ$ ), различия по амплитуде между зубцами R и S ( $daRS$ ) и пиковым значениям волн (S, R, P и Q). Все эти особенности проиллюстрированы на Рис. 6. Нормализацию вышеупомянутых признаков можно определить путем деления значения признака на среднее его значение за последние тридцать два такта. Подробное объяснение механизма извлечения представлено в **Приложении В**.
- **Интервалы RR и их нормализация.** Это наиболее часто используемые функции для целей классификации. Эти характеристики можно рассчитать по временному расстоянию между последовательными ударами. Если мы обозначим  $RR_p$  для текущего интервала RR, то  $RR_p - 1$  и  $RR_p + 1$  представляют собой интервалы до RR и после RR соответственно. Число интервалов RR, рассматриваемых в данной статье, равно трем. Следующие шесть особенностей представляют собой нормализованные интервалы RR:  
 Нормализованный  $RR_p = (RR_p) / \text{среднее значение последних тридцати двух интервалов RR}$ .

Нормализованный  $RR_p - 1 = (RR_p - 1) / \text{среднее значение последних тридцати двух интервалов RR}$ .

Нормализованный  $RR_p + 1 = (RR_p + 1) / \text{среднее значение последних тридцати двух интервалов RR}$ .

Нормализованный  $RR_{r1} = (RR_p - 1) / RR_p$ .

Нормализованный  $RR_{r2} = (RR_p + 1) / RR_p$ .

Нормализованный  $RR_{pt} = [RR_p - \text{среднее значение последних тридцати двух интервалов RR}] / \text{стандартное отклонение последних тридцати двух интервалов RR}$ .

## 2.6. ВЫБОР ФУНКЦИИ

Чтобы минимизировать вычислительные затраты введенной структуры без ущерба для производительности классификатора, количество функций для этапа обучения было доведено до десяти; это согласуется с выводами специалистов и предыдущими работами [27,41]. Принцип взаимного информационного ранжирования (Mutual Information Ranking, MIR) используется для выбора только наиболее значимых функций в соответствии с классом сердцебиений. С точки зрения выбора признаков, MIR признан впечатляющей стратегией, поскольку он может определять нелинейную релевантность между признаками данного вектора признаков или матрицы признаков. Оценки MIR (или значения MI), оцененные между метками классов и признаков, указывают на способность различать эти признаки:

$MI(\text{feature}, \text{class label}) = E(\text{feature}) - E(\text{class label} | \text{feature})$  (8)

где  $E$  (особенность) – энтропия функции;  $E(\text{метка класса} | \text{функции})$  – условная энтропия метки класса для данной функции, оценивающая неопределенность метки класса всякий раз, когда функция известна. В этом исследовании значения MI оцениваются на основе работы в [42] и [43] и с использованием функции Python под названием "mutual\_info\_classif".

## 2.7. КЛАССИФИКАЦИЯ СЕРДЦЕБИЕНИЙ

В этом исследовании RF-классификатор используется для задачи классификации сердцебиений. Это ансамбль Z-деревьев  $R1(I), R2(I), \dots, RZ(I)$ , где  $I = i_1, i_2, \dots, i_j$  — вектор входных данных j-мерности, а полученная группа генерирует Z – выводит  $O1 = R1(I), O2 = R2(I), \dots, Oz = RZ(I)$ .  $Oz$  — значение прогноза, сделанное деревом номер z. Окончательный прогноз  $O$  определяется путем агрегирования результатов всех случайных деревьев (голосование большинства). RF производит Z-количество DT из L-обученных выборок. Новая обучающая группа создается путем



проведения бутстреп-выборки для всех деревьев в лесу индивидуально. Эта группа используется для разработки DT без обрезки. В каждом разделении узла DT случайным образом выбирается лишь небольшое количество  $j$  признаков, а не все из них. Чтобы построить случайно созданный граф, эта процедура повторялась для генерации  $J$  из DT.

Ниже приводится краткое описание процесса обучения случайно созданного графа:

**Этап первый:** Из обучающей группы выбрать произвольную выборку.

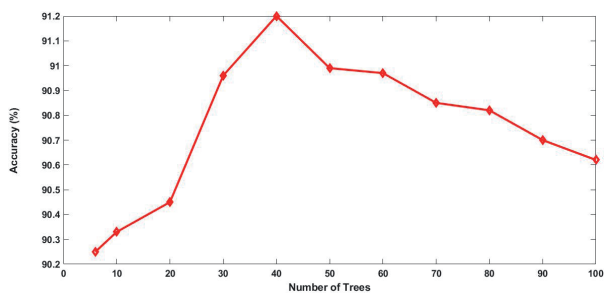
**Второй этап:** для каждой произвольной выборки создать дерево со следующим изменением: в каждом узле выбрать оптимальное разделение среди произвольно выбранной подгруппы входных параметров, которая является параметром настройки RF. Дерево создается полностью до тех пор, пока не останется потенциальных делений и оно не будет обрезано.

**Третий этап:** повторить первый и второй этапы, пока не будет создано  $Z$  таких деревьев.

Классификатор RF был создан с использованием библиотеки *Python Scikit-learn* [44]. Все переменные остались в конфигурации по умолчанию, кроме количества деревьев. Чтобы обнаружить наилучшее количество деревьев, обучение проводилось с увеличенным количеством деревьев с использованием десяти оптимальных функций и проверялось достижение ACC. На **Рис. 7** показаны результаты применения LOOCV на обучающем наборе. Понятно, что при числе деревьев, превышающем сорок, точность классификаторов не улучшается, а может даже ухудшаться.

### 2.8. ПОКАЗАТЕЛИ ОЦЕНКИ КЛАССИФИКАЦИИ

Измерение эффективности классификации и сравнение с современным уровнем техники проводилось путем определения следующих показателей: общая точность (ACC<sub>0</sub>), показатель F1, положительная прогнозируемость (PPr) и чувствительность (Sen).



**Рис. 7.** Зависимость точности от количества деревьев, полученных из LOOCV на обучающем наборе для RF-классификаторов.

$$F1 - Score = \frac{2 \cdot PPr \cdot Sen}{PPr + Sen} \quad (8)$$

$$PPr = \frac{TP}{TP + FP} \quad (9)$$

$$Sen = \frac{TP}{TP + FN} \quad (10)$$

$$ACC_0 = \frac{TP + TN}{TP + TN + FP + FN} \quad (11)$$

где TP – истинно положительный результат, TN – истинно отрицательный результат, FP – ложный положительный результат и FN – ложный отрицательный результат.

### 3. РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

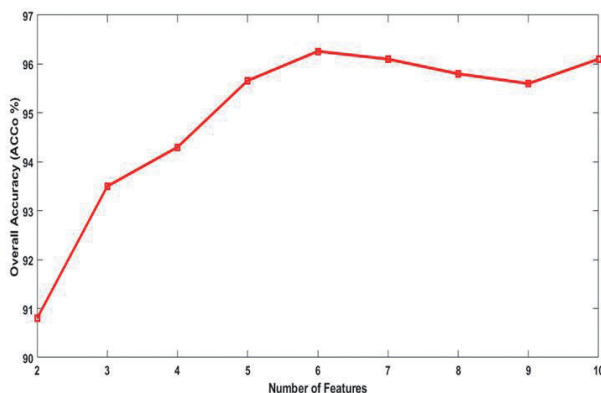
В этой статье проводится исследование использования радиочастотного классификатора для классификации сигналов ЭКГ путем использования оптимального выбора для извлеченных функций временной области (TD) и соблюдения стандартов АAMI и подхода между пациентами. В ходе этого исследования были изучены различные восемьдесят пять особенностей. Были реализованы все функции, предложенные в [37], которые считаются важной исследовательской работой в этом направлении. Кроме того, было использовано большинство функций, предложенных или основанных на [41] и [38]. В этом исследовании нормализация была достигнута путем деления характеристик на среднее значение последних сердечных сокращений. Этот процесс можно более практично проводить с помощью носимых и портативных медицинских устройств мониторинга в режиме реального времени. Дальнейшую нормализацию  $RRp - 1$  и  $RRp + 1$  проводили путем деления их на интервал  $RRp$ ; это было представлено как попытка учесть временную значимость между последовательными пиками  $R$ , что также рассматривается в [45]. Нормализованный  $RRp_{pt}$  рассчитывался как количественная мера стандартных отклонений для определенного интервала от среднего значения. **Таблица 4** иллюстрирует десять наиболее важных характеристик, оцениваемых по принципу MIR. Поскольку высокая доля вышеуказанных нормировок представлена в таблице наиболее значимых особенностей (как указано в Таблице 4), наши результаты показывают, что нормализованные интервалы  $RR$  предоставляют дополнительную информацию и более эффективны для классификации сигналов ЭКГ, чем другие классические интервалы  $RR$ . В

Таблица 4

Десять наиболее значимых характеристик, полученных с помощью MIR, с их значениями MI

Позиция по версии MIR	Имя функции	Значение MI
1	Нормализованный QRSm	0.19037819
2	Нормализованный QRSq	0.17508354
3	Нормализованный PPr	0.16311836
4	Нормализованный PPr2	0.14615831
5	QRSm	0.14439339
6	Первый аппроксимирующий коэффициент четвертой степени HBF (HBF1D4)	0.14157413
7	HBF1D4	0.14001679
8	QRSq	0.13641600
9	HBF1D4	0.13299729
10	Нормализованный RRr1	0.13147040

тестах была задействована только лучшая десятка с наиболее значимыми характеристиками. Результаты показывают, что шесть функций достаточны для достижения оптимальных результатов ( $ACC_o = 96.26\%$ ), как показано на **Рис. 8**. Это согласуется с полученными результатами в [38], которые показывают количество наиболее отличительных особенностей. Подмножество обычно невелико, и исключение менее значимых из них может эффективно повысить производительность классификатора. Помимо особенностей нормализованных интервалов RR, наибольшее количество из десяти наиболее значимых особенностей связано с длительностью комплекса QRS, рассчитанной при определенных величинах. Это указывает на то, что эти конкретные особенности являются более надежными идентификаторами формы QRS, чем контрольные точки, которые характеризуются вариациями небольшой амплитуды и которые часто нарушаются фоновым шумом. **Таблицы**



**Рис. 8.** Общая точность ( $ACC_o$  %) в зависимости от количества признаков RF-классификатора.

Таблица 5

Производительность предлагаемого метода с точки зрения матрицы путаницы с использованием тестового набора

		Прогноз		
		NB	SVEB	VEB
	NB	43494	653	69
	SVEB	367	1447	23
Истинный	VEB	377	15	2818
	FB	269	2	67
	QB	3	1	5

**5 и 6** демонстрируют эффективность предлагаемой структуры с точки зрения матрицы путаницы и различных показателей оценки соответственно. Поскольку неправильная классификация QB и FB не принимается во внимание, согласно рекомендациям ААММИ, они не рассматриваются в настоящем исследовании (как наиболее предшествующие работы), но все же участвуют в оценке. Примечательно, что показатель F1 дает более надежное представление о производительности классификаторов, особенно когда данные несбалансированы. В контексте сравнения с предыдущими работами предлагаемая структура отличается от подходов предыдущих исследований тем, что является комплексной и полностью автоматизированной. В отличие от предыдущих методов, которые не определяли автоматически местоположения пиков R (они работают только в том случае, если эти местоположения заранее определены в базе данных), предлагаемая структура представляет новый, простой и эффективный метод обнаружения этих мест на основе тригонометрии. В **Таблице 7** сравниваются результаты, полученные с помощью тестового набора, представленного в этой статье, с несколькими методологиями, представленными в литературе, которые достигли очень хороших показателей классификации сигналов ЭКГ и использовали рекомендации ААММИ и концепцию взаимодействия с пациентами. Показано, что общая точность, полученная в данной работе, является одной из лучших, достигаемых при

Таблица 6

Эффективность предлагаемого подхода с точки зрения Sen, PPr, F1-Score и  $ACC_o$  с использованием тестового набора.

	Sen	PPr	F1-Score	$ACC_o$
NB	98.36%	97.71%	98.04%	
SVEB	78.77%	68.32%	73.17%	<b>96.26%</b>
VEB	87.78%	94.50%	91.02%	

Таблица 7

## Сравнение эффективности предлагаемой структуры с современными методами

Источник	Извлеченные функции	Размер используемых функций	Классификатор	Класс F1 NB	Класс F1 SVEB	Класс F1 VEB	ACCо
[48]	Интервалы RR, морфология временной области и WT	101	Ансамбль SVM	97.59%	57.68%	92.37%	93.8%
[37]	Интервалы RR, морфология временной области, HOS и WT	45	Ансамбль SVM	97.04%	60.74%	94.29%	94.5%
[46]	GM, HOS и интервалы RR	33	Ансамбль BDT	97.89%	88.64%	85.82%	96.15%
[49]	Случайные проекции DCT, интервалы RR	33	SVM	96.88%	33.37%	77.29%	93.1%
[27]	Интервалы RR, морфология временной области и WT	8	Линейный дискриминант	96.48%	17.16%	83.89%	93%
[50]	HOS	4	Наивный Байес	65.6%	0.4%	84.7%	94.0%
	Интервалы RR, HOS, WT и морфология временной области	45	SVV и VTT	91.50%	54.45%	94.47%	88.69%
[33]	-	-	RCNN	93.93%	45.9%	83.89%	88.99%
Представленный доклад	Интервалы RR, морфология временной области и HBF	6	RF	<b>98.04%</b>	73.17%	91.02%	<b>96.26%</b>

использовании тех же показателей оценки (ACC = 96.26%), превосходит показатели, достигнутые в [46] (ACCо = 96.15%), что, по данным нашего исследования, самое высокое из предыдущих работ. Исследователи в [46] использовали подход Гауссовой смеси (GM) (который является дорогостоящим в вычислительном отношении) для получения различных характеристик, а также использовали метод упаковки дерева для целей классификации. Схема классификации RF может дать лучшие результаты, чем дерево упаковки, поскольку RF обеспечивает минимальную корреляцию между деревьями; это также было подчеркнуто в [47]. Предложенный метод в [46] использовал сто деревьев для получения результатов, тогда как результаты, представленные в этой статье, используют только сорок деревьев. Более того, предлагаемый метод является более общим и практичным в реальном мире, поскольку решает проблему WGN, которая может негативно повлиять на результаты классификации. Эта проблема не учитывается в [46] и большинстве предыдущих методов. Наилучшие результаты обнаружения SVEB были достигнуты в [46]; однако предложенная система превзошла результаты с точки зрения обнаружения VEB. Представленная система заняла второе место по выявлению VEB (F1-оценка = 91.02%) после первого ранга, достигнутого в [51] (F1-оценка = 94.47%). Тем не

менее, наша модель более эффективна с точки зрения сложности, простоты и пригодности для носимых и портативных устройств, чем модель, предложенная в [51], основанная на DNN. Выявление VEB особенно важно, поскольку его рецидив может быть использован как предиктор опаснейшего заболевания "сердечная недостаточность", часто приводящего к летальному исходу [14,15].

#### 4. ЗАКЛЮЧЕНИЕ

В этом исследовании была предложена комплексная, надежная, эффективная и полностью автоматизированная система классификации сигналов ЭКГ на основе цифровой обработки сигналов, инженерного выбора функций и радиочастотного классификатора. Оптимальное сочетание разработки функций и радиочастотного классификатора было введено для решения задачи классификации пульса для устройств с ограниченными ресурсами. Эффективность классификации сигналов ЭКГ была разумно рассчитана с учетом рекомендаций ААМ и индивидуального подхода к пациентам. Наиболее значимыми особенностями для задачи классификации ЭКГ были нормализованные характеристики, связанные с интервалами RR и продолжительностью комплекса QRS, рассчитанные при

определенных величинах. Оптимальные результаты были достигнуты с использованием шести лучших наиболее значимых функций и сорока деревьев решений RF-классификатора. Оценка была основана на MTT-BIN-ARR-DB. Сравнивая предложенную в этой статье схему с опубликованными в литературе, можно отметить, что наши результаты являются одними из лучших результатов, достигнутых на сегодняшний день. Результаты не только подтвердили, что RF-классификатор является превосходным подходом для классификации сигналов ЭКГ, но также сравнительно небольшое количество функций и деревьев адекватно для достижения или превосходства литературных показателей.

#### Приложение А Настройка $Y_1$ и $Y_2$ для получения наименьшей ошибки обнаружения

	$Y_2$				
$Y_1$	0.00006	0.00008	0.00018	0.00012	0.00014
0.2	0.350	0.281	0.262	0.251	0.624
0.3	0.235	0.218	0.216	0.207	0.218
0.4	0.198	0.191	0.190	<b>0.185</b>	0.193
0.5	0.198	0.196	0.195	0.193	0.200
0.6	0.200	0.195	0.208	0.203	0.220

#### Приложение Б

##### Извлечение ключевых точек сигнала ЭКГ

Наивысшая абсолютная амплитуда сигнала в период 100 мс до и после пика R назначается в качестве опорной базовой точки. Для получения ключевых точек ЭКГ были проведены следующие процедуры: (обратите внимание, что пики R обнаруживаются и извлекаются, как описано в разделе 2.3).

1. Первоначально предположим, что пики P, Q и S = 0 (соответствующих волн не существует).
2. Вернуться назад перед QRSmp и определить точки перехода следующим образом:
  - i. Установить  $[QRSmp/2p]$  = первую позицию, в которой значение сигнала ниже QRSmp/2.
  - ii. Установить  $[QRSmp/4p]$  = первую позицию, в которой значение сигнала ниже QRSmp/4.
  - iii. Установить  $Q_{peak}$  = амплитуду первой отрицательной точки перехода.
  - iv. Если первая точка перехода ( $t_{rp}$ )  $\geq 0$ , то указать эту точку как QRSb, установить  $S_{peak} = QRSmp$  и  $Q_{peak} = 0$ .
  - v. Если QRSmp  $> 0$ , а второй  $t_{rp} < 0$  и  $Q_{peak} = 0$ , то указать амплитуду этой точки как  $Q_{peak}$ .

- vi. Если  $Q_{peak} \neq 0$  и сигнал проходит через 0, то установить QRSb = первую точку, которая равна  $\geq 0$ .
- vii. Если второй  $t_{rp} \geq 0$ , а QRSb еще не определен, то указать эту точку как QRSb.

3. Продолжить после QRSmp и определить точки перехода следующим образом:

- i. Установить  $[QRSmp/2n]$  = первую позицию, в которой значение сигнала ниже QRSmp/2.
- ii. Установить  $[QRSmp/4n]$  = первую позицию, в которой значение сигнала ниже QRSmp/4.
- iii. Установить  $S_{peak}$  = амплитуде первого  $t_{rp}$ , то есть  $< 0$ .
- iv. Если  $S_{peak} \neq 0$  и сигнал проходит через 0, то установить QRSe = первую точку, которая  $\geq 0$ .

4. Оценить максимальную амплитуду сигнала в период от 236 мс до 70 мс до QRSb (MApriQRSb[236,70]). Если эта амплитуда  $> 3 \cdot STD$  сигнала за период 70 мс до рассматриваемого периода и ее местоположение соответствует  $t_{rp}$  сигнала, то устанавливаем  $P_{peak} = MApriQRSb[236,70]$ .

Примечание: QRSb – начало QRS, QRSe – окончание QRS, STD – стандартное отклонение.

#### ЛИТЕРАТУРА

1. Xu J, Xiao W, Liang X, Shi L, Zhang P, Wang Y, Wang Y, Yang H. A meta-analysis on the risk factors adjusted association between cardiovascular disease and COVID-19 severity. *BMC Public Health*, 2021, 21(1):1533; doi: 10.1186/s12889-021-11051-w.
2. Tyagi A, Mehra R. 2021. Intellectual heartbeats classification model for diagnosis of heart disease from ECG signal using hybrid convolutional neural network with GOA. *SN Applied Sciences*, 2021, 3(2):1-14.
3. Pal HS, Kumar A, Vishwakarma A, Singh GK, Lee HN. An effective ECG signal compression algorithm with self-controlled reconstruction quality. *Computer Methods in Biomechanics and Biomedical Engineering*, 2023, pp. 1-11. DOI: 10.1080/10255842.2023.2206933.
4. Ali AM, Ahmed AF, Najim AH. 2020, November. Efficient and Effective Scheme for ECG Compression. *Proc. 2nd Annual International Conference on Information and Sciences (AiCIS)*, 2020, pp. 91-94. DOI: 10.1109/AiCIS51645.2020.00024.
5. Martínez JP, Almeida R, Olmos S, Rocha AP, Laguna P. A wavelet-based ECG delineator: evaluation on standard databases. *IEEE Transactions on biomedical engineering*, 2004, 51(4):570-581.
6. Sharma M, Dhiman HS, Acharya UR. 2021. Automatic identification of insomnia using optimal antisymmetric biorthogonal wavelet filter bank with ECG signals. *Computers in Biology and Medicine*, 2021, 131:104246.

7. Hussain I, Park SJ. Big-Ecg: Cardiographic Predictive Cyber-Physical System for Stroke Management. *IEEE Access*, 2021, 9:123146-123164.
8. Chandra S, Sharma A, Singh GK. A comparative analysis of performance of several wavelet based ECG data compression methodologies. *IRBM*, 2021, 42(4):227-244. DOI: 10.1016/j.irbm.2020.05.004.
9. Yan Z, Zhou J, Wong WF. 2021. Energy efficient ECG classification with spiking neural network. *Biomedical Signal Processing and Control*, 2021, 63:102170.
10. Saini I, Singh D, Khosla A. 2013. QRS detection using K-Nearest Neighbor algorithm (KNN) and evaluation on standard ECG databases. *Journal of advanced research*, 2013, 4(4):331-344.
11. Martis RJ, Acharya UR, Mandana KM, Ray AK, Chakraborty C. Cardiac decision making using higher order spectra. *Biomedical Signal Processing and Control*, 2013, 8(2):193-203.
12. Kim H, Kim S, Van Helleputte N, Artes A, Konijnenburg M, Huisken J, Van Hoof C, Yazicioglu RF. A configurable and low-power mixed signal SoC for portable ECG monitoring applications. *IEEE transactions on biomedical circuits and systems*, 2013, 8(2):257-267.
13. Scirè A, Tropeano F, Anagnostopoulos A, Chatzigiannakis I. Fog-computing-based heartbeat detection and arrhythmia classification using machine learning. *Algorithms*, 2019, 12(2):32.
14. Acharya T, Tringali S, Bhullar M, Nalbandyan M, Ilineni VK, Carbajal E, Deedwania P. Frequent atrial premature complexes and their association with risk of atrial fibrillation. *The American journal of cardiology*, 2015, 116(12):1852-1857.
15. Liaqat S, Dashtipour K, Zahid A, Assaleh K, Arshad K, Ramzan N. Detection of atrial fibrillation using a machine learning approach. *Information*, 2020, 11(12):549. DOI:10.3390/info11120549.
16. Pyakillya B, Kazachenko N, Mikhailovsky N. Deep learning for ECG classification. *Journal of Physics Conference Series*, 2013, 913(1):012004. DOI: 10.1088/1742-6596/913/1/012004.
17. Hannun AY, Rajpurkar P, Haghpanahi M, Tison GH, Bourn C, Turakhia MP, Ng AY. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nature medicine*, 2019, 25(1):65-69.
18. Murat F, Yildirim O, Talo M, Baloglu UB, Demir Y, Acharya UR. Application of deep learning techniques for heartbeats detection using ECG signals-analysis and review. *Computers in biology and medicine*, 2020, 120:103726.
19. Luz EJDS, Schwartz WR, Cámara-Chávez G, Menotti D. 2016. ECG-based heartbeat classification for arrhythmia detection: A survey. *Computer methods and programs in biomedicine*, 2016, 127:144-164.
20. Ani R, Krishna S, Anju N, Aslam MS, Deepa OS. Iot based patient monitoring and diagnostic prediction tool using ensemble classifier. *Proc. International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 2017, pp. 1588-1593.
21. Saffari A, Leistner C, Santner J, Godec M, Bischof H. On-line random forests. *Proc. IEEE 12th international conference on computer vision workshops*, 2009, pp. 1393-1400.
22. Lempitsky V, Verhoek M, Noble JA, Blake A. Random forest classification for automatic delineation of myocardium in real-time 3D echocardiography. *Proc. International Conference on Functional Imaging and Modeling of the Heart*, 2009, pp. 447-456. Springer, Berlin, Heidelberg.
23. Vamsi IV, Abhinav N, Verma AK, Radhika S. Random forest based real time fault monitoring system for industries. *Proc. 4th International Conference on Computing Communication and Automation (ICCCA)*, 2018, pp. 1-6.
24. Zhao X, Song Z, Guo J, Zhao Y, Zheng F. Real-time hand gesture detection and recognition by random forest. *Proc. Int. Conf. "Communications in computer and information processing" (ICCIP)*, 2012, 289:747-755. Springer, Berlin, Heidelberg.
25. Li J, Zhong PA, Yang M, Zhu F, Chen J, Liu W, Xu S. Intelligent identification of effective reservoirs based on the random forest classification model. *Journal of Hydrology*, 2020, 591:125324.
26. Emanet N. ECG beat classification by using discrete wavelet transform and Random Forest algorithm. *Proc. Fifth International Conference on Soft Computing, Computing with Words and Perceptions in System Analysis, Decision and Control*, 2009, pp. 1-4.
27. Llamedo M, Martínez JP. 2010. Heartbeat classification using feature selection driven by database generalization criteria. *IEEE Transactions on Biomedical Engineering*, 2010, 58(3):616-625.
28. Alickovic E, Subasi A. Medical decision support system for diagnosis of heart arrhythmia using DWT and random forests classifier. *Journal of medical systems*, 2016, 40(4):108.
29. Gutierrez-Gnecchi JA, Morfin-Magaña R, del Carmen A, Tellez-Anguiano DLE, Reyes-Archundia E, Díaz OH. Cardiac Arrhythmia Classification Using a Combination of Quadratic Spline-Based Wavelet Transform and Artificial Neural Classification Network. *Proceedings 2nd International Work-Conference on Bioinformatics and Biomedical Engineering (IWBBIO)*, 2014, Vols 1, pp.1743-1754.
30. Kumar RG, Kumaraswamy YS. Investigating cardiac arrhythmia in ECG using random forest classification. *Int. J. Comput. Appl*, 2012, 37(4):31-34.

31. Park J, Lee S, Kang K. Arrhythmia detection using amplitude difference features based on random forest. *Proc. 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2015, pp. 5191-5194.
32. Aravind S, Sanjay M. ECG Classification and Arrhythmia Detection Using Wavelet Transform and Convolutional Neural Network. *Proc. International Conference on Communication, Control and Information Sciences (ICCISc)*, 2021, Vol. 1, pp. 1-5.
33. Li Y, Qian R, Li K. Inter-patient arrhythmia classification with improved deep residual convolutional neural network. *Computer Methods and Programs in Biomedicine*, 2022, 214:106582.
34. <https://physionet.org/content/mitdb/1.0.0/>.
35. Wang T, Lu C, Sun Y, Yang M, Liu C, Ou C. 2021. Automatic ECG Classification Using Continuous Wavelet Transform and Convolutional Neural Network. *Entropy*, 2021, 23(1):119.
36. De Chazal P, O'Dwyer M, Reilly RB. Automatic classification of heartbeats using ECG morphology and heartbeat interval features. *IEEE transactions on biomedical engineering*, 2004, 51(7):1196-1206.
37. Mondéjar-Guerra V, Novo J, Rouco J, Penedo MG, Ortega M. Heartbeat classification fusing temporal and morphological information of ECGs via ensemble of classifiers. *Biomedical Signal Processing and Control*, 2019, 47:41-48.
38. De Lannoy G, François D, Delbeke J, Verleysen M. Weighted SVMs and feature relevance assessment in supervised heart beat classification. *Proc. International Joint Conference on Biomedical Engineering Systems and Technologies*, 2010, pp. 212-223. Springer, Berlin, Heidelberg.
39. Osowski S, Linh TH. ECG beat recognition using fuzzy hybrid neural network. *IEEE Transactions on Biomedical Engineering*, 2001, 48(11):1265-1271.
40. Ince T, Kiranyaz S, Gabbouj M. A generic and robust system for automated patient-specific classification of ECG signals. *IEEE Transactions on Biomedical Engineering*, 2009, 56(5):1415-1426.
41. Mar T, Zaunseder S, Martínez JP, Llamedo M, Poll R. Optimization of ECG classification by means of feature selection. *IEEE transactions on Biomedical Engineering*, 2011, 58(8):2168-2177.
42. Kraskov A, Stögbauer H, Grassberger P. Estimating mutual information. *Physical review E*, 2004, 69(6):066138.
43. Ross BC. Mutual information between discrete and continuous data sets. *PloS one*, 2014, 9(2):e87357.
44. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J. Scikit-learn: Machine learning in Python. *The Journal of machine Learning research*, 2011, 12:2825-2830.
45. Tsipouras MG, Fotiadis DI, Sideris D. An arrhythmia classification system based on the RR-interval signal. *Artificial intelligence in medicine*, 2005, 33(3):237-250.
46. Afkhami RG, Azarnia G, Tinati MA. Cardiac arrhythmia classification using statistical and mixture modeling features of ECG signals. *Pattern Recognition Letters*, 2016, 70:45-51.
47. Banfield RE, Hall LO, Bowyer KW, Kegelmeyer WP. A comparison of decision tree ensemble creation techniques. *IEEE transactions on pattern analysis and machine intelligence*, 2006, 29(1):173-180.
48. Huang H, Liu J, Zhu Q, Wang R, Hu G. A new hierarchical method for inter-patient heartbeat classification using random projections and RR intervals. *Biomedical engineering online*, 2014, 13(1):1-26.
49. Chen S, Hua W, Li Z, Li J, Gao X. Heartbeat classification using projected and dynamic features of ECG signal. *Biomedical Signal Processing and Control*, 2017, 31:165-173.
50. Marinho LB, de MM Nascimento N, Souza JWM, Gurgel MV, Reboucas Filho PP, de Albuquerque VHC. A novel electrocardiogram feature extraction approach for cardiac arrhythmia classification. *Future Generation Computer Systems*, 2019, 97:564-577.
51. Li Y, He Z, Wang H, Li B, Li F, Gao Y, Ye X. CraftNet: a deep learning ensemble to diagnose cardiovascular diseases. *Biomedical Signal Processing and Control*, 2021, 62:102091.

**Анас Фуад Ахмед**

доцент

Иракский университет

**Багдад 6029, Ирак**

**E-mail: anasfuad33eng@yahoo.com**

**Хадил Н. Абдуллах**

доктор наук, проф.

Технологический университет

**Багдад 19006, Ирак**

**E-mail: 30002@uotechnology.edu.iq**

**Бараа М. Албакер**

доцент

Иракский университет

**Багдад 6029, Ирак**

**E-mail: baraamalbaker@ymail.com.**