# Arrhythmia Diagnosis by Using Level-Crossing ECG Sampling and Sub-Bands Features Extraction for Mobile Healthcare

**Saeed Mian Qaisar [1],* and Syed Fawad Hussain [2]**

[1] College of Engineering, Effat University, Jeddah 22332, Saudi Arabia
[2] Machine Learning and Data Science Lab, Ghulam Ishaq Khan Institute of Engineering Sciences and Technology, Topi 23460, Pakistan; fawadhussain@giki.edu.pk
* Correspondence: sqaisar@effatuniversity.edu.sa; Tel.: +966-1221-37849

**Abstract:** Mobile healthcare is an emerging technique for clinical applications. It is usually based on cloud-connected biomedical implants. In this context, a novel solution is presented for the detection of arrhythmia by using electrocardiogram (ECG) signals. The aim is to achieve an effective solution by using real-time compression, efficient signal processing, and data transmission. The system utilizes level-crossing-based ECG signal sampling, adaptive-rate denoising, and wavelet-based sub-band decomposition. Statistical features are extracted from the sub-bands and used for automated arrhythmia classification. The performance of the system was studied by using five classes of arrhythmia, obtained from the MIT-BIH dataset. Experimental results showed a three-fold decrease in the number of collected samples compared to conventional counterparts. This resulted in a significant reduction of the computational cost of the post denoising, features extraction, and classification. Moreover, a seven-fold reduction was achieved in the amount of data that needed to be transmitted to the cloud. This resulted in a notable reduction in the transmitter power consumption, bandwidth usage, and cloud application processing load. Finally, the performance of the system was also assessed in terms of the arrhythmia classification, achieving an accuracy of 97%.

**Keywords:** electrocardiogram (ECG); compression; arrhythmia classification; level-crossing sampling; adaptive-rate processing; hysteresis; machine learning; mobile healthcare; wavelet decomposition; sub-bands features extraction

## 1. Introduction

An electrocardiogram (ECG) signal possesses critical information about cardiac functionality [1]. An abnormality of cardiac rhythm is a sign of certain diseases that can be diagnosed by an effective analysis of an ECG [2]. Heart diseases are one of the major threats to human life [3], and a timely diagnosis can lead to better measures. To do this, individual heartbeats are analyzed by exploiting their frequency content and morphological patterns for the automatic diagnosis of ECG arrhythmia [4–8]. Interferences and physiological artifacts can modify the ECG signal, thus decreasing the effectiveness of the automatic diagnosis mechanism. To overcome these deficiencies, numerous signal processing techniques have been used including eigenvalue composition [9], extended Kalman filtering [10], and Fourier decomposition [11]. The denoised ECG signals are processed and analyzed to extract classifiable features that can assist in automatic diagnosis. Some broadly used features extraction techniques are wavelet packet decomposition (WPD) [4], wavelet-based kernel principle component analysis (wkPCA) [4], wavelet packet entropy (WPE) [5], discrete wavelet transform (DWT) [6,12,13], temporal and morphological features [12], principle component analysis (PCA) [13,14], bispectrum [14], and Hermite function coefficient and temporal features [15].

The extracted features are employed for the automatic diagnosis of cardiac diseases. Supervised machine learning techniques are quite popular in these applications. In this regard, numerous approaches have been presented for ECG signal classification including those of the back propagation neural network (BNN) [4], random forest (RF) [5], probabilistic neural network (PNN) [6], support vector machine (SVM) [12], support vector machine (SVM)-radial basis function (RBF) [13,14], and optimized block-based neural network (OBNN) [15].

A timely detection of arrhythmia conditions allows for effective cardiac failure treatment. Therefore, patients with cardiac problems require continuous monitoring. In this context, wearable ECG sensors have been used [16,17]. These are usually linked to the cloud, and their data are logged and further processed via cloud-based applications. The decision support outcome is then shared with a monitoring center [17–20]. Such an approach allows healthcare specialists to make timely decisions and offer emergency care for patients with chronic disorders.

The design of ECG wearable devices is challenging because of strict constraints on size, weight, and power consumption. Self-powered wireless sensors are preferred because they enable the acquisition of ECG signals without causing too much uneasiness to the patient. A precise diagnosis requires the fine-grained recording and analysis of multichannel ECG signals. Consequently, the data dimension is exponentially augmented [21]. The transmission, analysis, and storage of such a big amount of data are not efficient. The substantial saving of power in a wireless ECG device is only possible by minimizing the activity of acquisition, processing, and wireless data transmission. In this context, studies have been carried out on ECG signal compression [20–22], non-uniform sampling [16,23,24], and adaptive-rate processing [25].

Conventional ECG processing solutions are based on the Nyquist sampling theory. Systems are time-invariant, which results in a worst-case parameterization [25–29]. A system's computational load and processing activity remain fixed regardless of the incoming ECG signal time-varying nature. As a result, the system is constrained. It captures and processes redundant samples that increase the system's overall computational load, power consumption, processing, and transmission activities [27,29]. These shortfalls can be compensated for to a certain extent by using level-crossing analog-to-digital converters (LCADCs) [16,24,30–35]. These converters adapt the system acquisition and processing rates according to the incoming signal's temporal variations, a process that renders a significant computational efficiency of the suggested approach compared to traditional ones. In [27,29], Qaisar et al. reported more than one order of magnitude gain in terms of computational outperformance compared to the classical counterparts.

This work contributes to the development of efficient automatic arrhythmia diagnosis in the mobile healthcare framework. The aim is to realize an effective solution by achieving real-time compression and efficient ECG signal processing and transmission. In a continuation of previous works, such as [16,24,28,30,31], level-crossing sampling and adaptive-rate processing techniques are proposed in this paper. The realization of such is achieved by using an efficient assembly of the LCADC, robust classifiers, adaptive-rate signal processing, and features extraction modules. The use of level-crossing sampling and adaptive-rate processing for arrhythmia detection is an original concept thatis explored in this paper.

## 2. Materials and Methods

The proposed framework is shown in Figure 1, where the dotted blue-colored border (' . . . .') encloses blocks embedded in the ECG real-time wearable processing device. The cloud processing module is enclosed in a green-colored dashed line ('---').
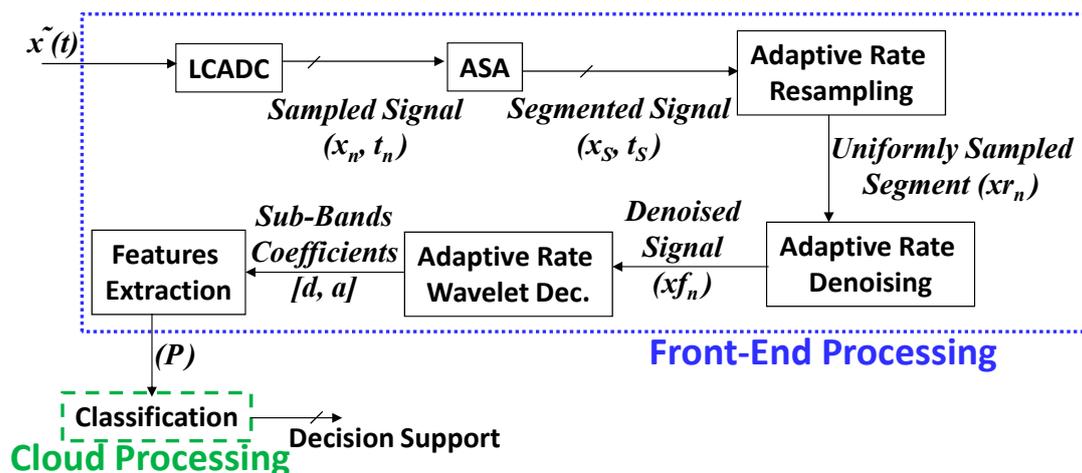
**Figure 1.** The proposed system block diagram. LCADC: level-crossing analog-to-digital converters; ASA: activity selection algorithm.

## 2.1. Dataset

The performance of the designed solution was studied by using the MIT-BIH arrhythmia database (ECG arrhythmia time series are available under (https://physionet.org/content/mitdb/1.0.0/)) [36]. A set of twelve 30-min ECG recordings containing clinically important arrhythmias was employed. Each channel was band-limited between [0.5, 60] Hz by using an analog antialiasing filter and was then acquired with an 11-bit resolution analog-to-digital converter (ADC). The sampling rate was 360 Hz. Experienced cardiologists labeled all acquired heartbeats. To extract individual heartbeats, each considered ECG signal wassegmented for a time length of 0.9 s. Five different arrhythmias classes were considered: normal signals (N), right bundle branch block (RBBB), left bundle branch block (LBBB), atrial premature contraction (APC) and premature ventricular contraction (PVC). To achieve equal representation, 300 instances were considered for each class. Therefore, in total, there were 1500 instances, each belonging to one of the five considered classes. To avoid any bias, heartbeats related to each class were collected from various records (clear from Table 1).

**Table 1.** MIT-BIH arrhythmia records used in this study. RBBB: right bundle branch block; LBBB: left bundle branch block; APC: atrial premature contraction; PVC: premature ventricular contraction.

| CLASS | Records | Number of Beats Used |
|---|---|---|
| Normal | 116, 119, 209 | 300 |
| RBBB | 118, 124, 212 | 300 |
| LBBB | 109, 111, 214 | 300 |
| APC | 118, 200, 209 | 300 |
| PVC | 119, 200, 233 | 300 |

## 2.2. Level-Crossing A/D Converter (LCADC)

To evaluate the LCADC, the sampled version of a band-limited signal, given by $y(t_n)$ and obtained from the MIT-BIH dataset, was reconstructed. If $\widetilde{x}(t)$ is the reconstructed quasi-analog version of $y(t_n)$, then the relationship between $\widetilde{x}(t)$ and $y(t_n)$ can be mathematically presented by using Equation (1), where $U$ is the up-sampling factor. The up-sampling is realized by using the standard cascaded cubic spline interpolators and anti-imaging filters [37].

$$\widetilde{x}(t) = y\left(t_{\frac{n}{U}}\right). \tag{1}$$

The band-limited signal $\widetilde{x}(t)$ is digitized with a LCADC. The frequency content of $\widetilde{x}(t)$ is limited to [0.5, 60] Hz [9,36]. The LCADC is designed on the basis of level-crossing sampling (LCS) [28,38].

In this case, a sample is only acquired when $\widetilde{x}(t)$ crosses one of the predefined thresholds. For a given LCADC amplitude dynamic, $\Delta V$, and resolution, $M$, the sampling frequency is piloted by the signal. Samples are irregularly spaced in time, and the count of samples is proportional to the slope of $\widetilde{x}(t)$ [31]. Sample amplitudes are equal to the predefined thresholds. The sampling instants are defined by Equation (2), and the process is shown in Figure 2 [28,38], where $t_n$ is the present sampling instant, $t_{n-1}$, is the preceding one and the time step between the present and the preceding sampling instants is $dt_n$. LCADC parameters are selected according to the approach described in [35]. This approach is based on the uniform-quantization scheme. Therefore, its quantum, $q$, can be calculated as: $q = \frac{\Delta V}{2^{M-1}}$ [38].

$$t_n = t_{n-1} + dt_n. \tag{2}$$

The phenomenon of hysteresis is also embedded in the LCADC [29]. A new sample is only acquired when there is a difference of $q$ with respect to the preceding sample amplitude. The process is shown in Figure 2 and can be mathematically expressed as: $x_n = x_{n-1} \pm q$. It improves LCADC efficiency in terms of real-time compression [29]. The QRS complexes of heartbeats contain the most significant arrhythmia-related information [16,24]. The LCADC acquires the relevant ECG information, QRS complexes, at adaptive-rates while avoiding the remaining low amplitude components [24]. Therefore, it collects a reduced number of samples while comparing them with the classical counterparts [24].
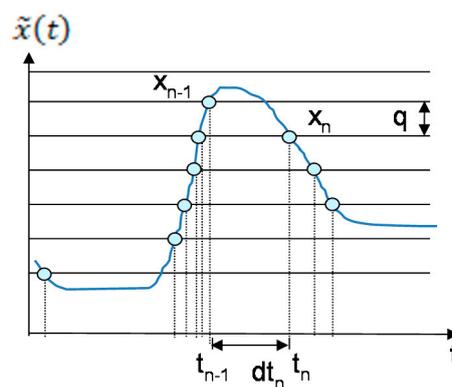


**Figure 2.** The principle of a level-crossing ADC with hysteresis.

The working principle of an LCADC is different from conventional ADCs [38]. While considering the ideal case, for conventional ADCs, the sampling instants are accurately known. However, the amplitudes of samples are quantized [38]. Quantization is the only source of error, and it depends on the selected $\Delta V$ and $M$ [35]. It is assessed in terms of the signal-to-noise ratio (SNR) [38]. The SNR is computable as: $SNR_{dB} = 6.02M + 1.76$. It presents an ideal ADC SNR for a full-scale monotonous sinusoidal input and describes its dependency on $M$. On the contrary, the amplitudes of samples are ideally known for an LCADC. However, the instants of these samples are quantized according to the operating frequency, $F_{Timer}$, of the timer circuit thatis used to record these instants [28]. The SNR of an ideal LCADC can be calculated by using Equation (3) [38], where $f_{sig}$ is the frequency of the full-scale sinusoid used to evaluate the LCADC. Equation (3) shows that the SNR of an ideal LCADC is independent of $M$ and is a function of $f_{sig}$ and $T_{Timer} = \frac{1}{F_{Timer}}$. A 6.02 dB improvement in the value of the SNR is achievable by halving $T_{timer}$ [28]. In this study, a 21-bit resolution timer wasused with $F_{Timer} = 1$ MHz. These parameters allowed us to properly record a heartbeat without timer overflow and resulted in an ideal LCADC SNR of 73.25 dB. According to [4–6,12–15], an11-bit ADC resolution is appropriate and results in a precise, computer-based arrhythmia diagnosis. For the selected timer parameters, the obtained LCADC SNR wasequal to the theoretical SNR of an 11.9-bit classical ADC. This justified the selected system parameters for the considered application.

$$SNR_{dB} = -11.19 - 20log\left(f_{sig} \cdot T_{timer}\right) \tag{3}$$

### 2.3. Activity Selection Algorithm (ASA)

The activity selection algorithm (ASA) segments LCADC output [27,28]. Thisalgorithmemploys sampling process non-uniformity for selecting the active parts of the signal while avoiding the redundant baseline [16,24,28]. The principle is clear from the algorithmic state machine (ASM) chart shown in Figure 3. In Figure 3, $T_0 = \frac{1}{f_{min}}$ is the fundamental period of $\widetilde{x}(t)$, and $f_{min}$ is the lowest frequency component whose value is equal to 0.5 Hz [9,36]. The level-crossing concept based sampled signal active parts are identified by using values of $T_0$ and $dt_n$. The condition $dt_n \leq T_0$ is selected to respect the Nyquist criterion for $f_{min}$. $L^i$ is the length of the $i^{th}$ selected segment $W^i$. $N^i$ is the number of samples for $W^i$. $L_{ref}$ is the superior bound on $L^i$, and its choice depends on the system parameters and the characteristics of the intended signal [28,38]. For this study $L_{ref}$ = 1-swas selected. At the beginning of each iteration, '$i$' is incremented, and $N^i$ and $L^i$ are initialized to zero.

The traditional windowing functions do not provide interesting features of the ASA [27,29]. The ASA allows for the selection of the signal active portions while avoiding the redundant, unwanted ones [16,24,28]. In addition, the length of the window-function is automatically adjusted according to the temporal variations of the signal. This process avoids signal truncation, and, therefore, segmentation can be performed by using adaptive length rectangular windows. This avoids the use of arithmetically complex smoothening-window functions and creates an effective solution of the spectral leakage phenomenon [27].
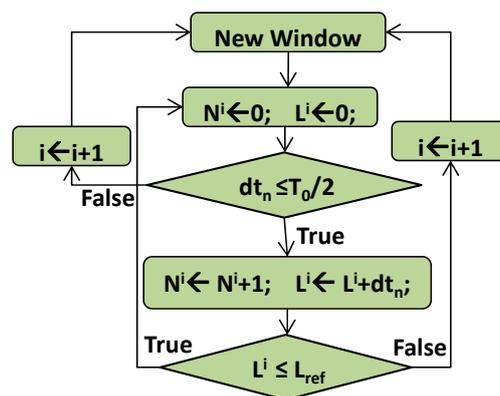


**Figure 3.** The algorithmic state machine (ASM) chart of activity selection algorithm.

### 2.4. Adaptive-Rate Resampling

For a given resolution $M$, the LCADC sampling frequency organizes by following the temporal variations of $\widetilde{x}(t)$. The maximum sampling frequency, $Fs_{max}$, of a uniform quantization-based LCADC is defined by Equation (4) [28,38], where $f_{max}$ is the $\widetilde{x}(t)$ bandwidth. $A_{in}$ is the input signal amplitude.

$$Fs_{max} = 2 \cdot f_{max} \left( 2^M - 2 \right) \cdot \frac{A_{in}}{\Delta V} \tag{4}$$

The sampling frequency for $W^i$ can be calculated as: $Fs^i = \frac{N^i}{L^i}$. To benefit from the established signal processing techniques, $W^i$ is uniformly resampled by using simplified linear interpolation (SLI) [29], which modifies the resampled signal compared to the original; this variation depends on $M$, $q$, and the employed interpolator [39]. For SLI, the superior bound of error per resampled observation is $\frac{q}{2}$ [39].

The LCADC focuses on the active signal parts. Nevertheless, one LCADC shortfall is that the active signal parts can be digitized at superior rates compared to conventional digitization approaches [27,29]. The ASA overcomes this shortfall by examining features of $W^i$ and thenadjusting the system parameters accordingly [29]. In this way, the resampling frequency, $Frs^i$, and the arithmetic complexity of the post-processing modules are adjusted by following the $\widetilde{x}(t)$ temporal disparities [29].

## 2.5. Adaptive-Rate Denoising

A band-pass finite impulse response (FIR) filters bank is designed for the effective online diminishing of unwanted noise from ECG signals [9]. The FIR filtering process can be mathematically described by using Equation (5), where $x_n$ is the incoming signal, $xf_n$ is the filtered signal, and $h_k$ are the coefficients of the $K^{th}$ order FIR filter.

$$xf_n = \sum_{k=0}^{K-1} h_k \times x_{n-k}. \tag{5}$$

Here, the filters bank was designed for the cut-off frequencies of [Fc$_{min}$, Fc$_{max}$] Hz. Each filter was implemented for a different sampling frequency that was chosen from the set *Fref* (cf. Equation (6)). The upper bound on *Fref* was selected as $F_r$ and, to assure a proper digital filtering operation, the lower bound on *Fref* was chosen as $Fs_{min} \geq 2 \cdot F_{Cmax}$ [29]. $F_r$ was the sampling frequency such that its value remained greater than and closer to $F_{Nyq} = 2 \cdot F_{Cmax}$. $Q$ is the length of *Fref*, and its value is always chosen as a binary-weighted. In Equation (6), $\Delta$ is a unique offset and can be computed as: $\Delta = \frac{F_r - Fs_{min}}{Q-1}$.

$$Fref = \{Fs_{min}, Fs_{min} + \Delta, \dots, Fs_{min} + (Q-1)\Delta = F_r\} \tag{6}$$

The ASA examines the properties of $W^i$ and uses them for adjusting the denoising parameters such as the resampling frequency, $Frs^i$, and the filter order, $K^i$. The method of choosing $Frs^i$ and keeping it aligned with $Fref_C$ is shown in Figure 4, which shows that a suitable filter, from the reference bank, is selected for each $W^i$. Let $h_{c_k}$ be the selected filter for $W^i$ that is sampled at $Fref_C$. Then, this selection can be made on the basis of *Fref* and $Fs^i$. For proper denoising, $Frs^i = Fref_C$ must bechosen [29].
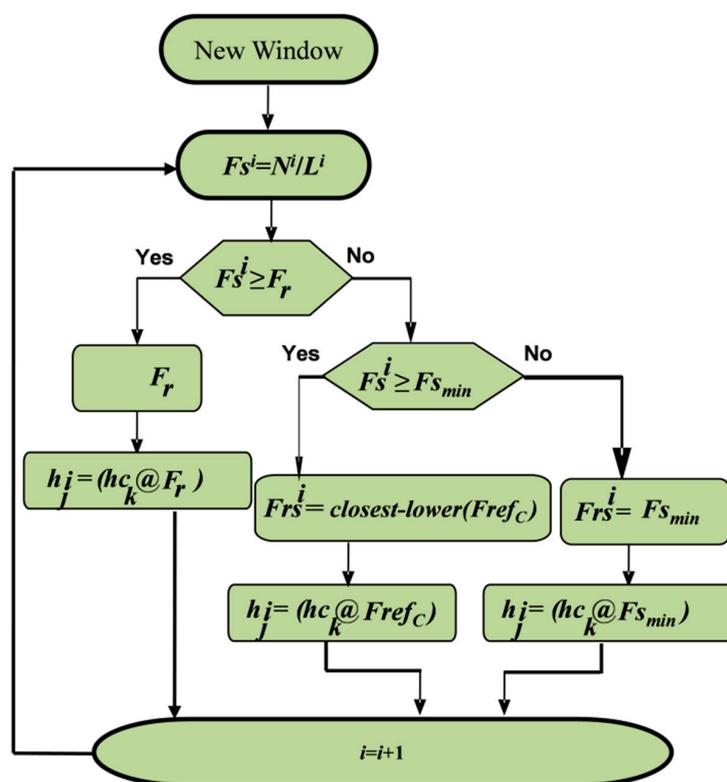


**Figure 4.** The process of selecting the resampling frequency and the filter coefficients from the reference filter bank for the $i^{th}$ selected segment.

## 2.6. Features Extraction

### 2.6.1. Adaptive-Rate Discrete Wavelet Transform (ARDWT)

The wavelet transform (WT) is frequently used for the multi-resolution time-frequency analysis of non-stationary ECG like signals [4,40]. This transform can be mathematically expressed by Equation (7), where $s$ and $u$, respectively, represent the dilation and the translation parameters.

$$W_x^\psi(u,s) = \frac{1}{\sqrt{S}} \int_{-\infty}^{+\infty} x(t)\psi * \left(\frac{(t-u)}{s}\right) dt \tag{7}$$

A discrete-time wavelet transform (DWT) is used for analyzing digital signals. A translation-dilation representation is attained by employing digital filters. In this case, a denoised segment, $W^i$, is passed through the Daubechies algorithm-based wavelet decomposition, which consists of half-band high-pass and low-pass filters. This allows for the computation of approximation, $a_m^i$, and detail, $d_m^i$, coefficients at each level of decomposition. The mathematical processes of computing $a_m^i$ and $d_m^i$ are, respectively, expressed by Equations (8) and (9), where $m$ represents the level of decomposition. In this study, a fourth level of decomposition was performed, i.e., $m \in \{1,2,3,4\}$. $g_{2n-k}$ and $h_{2n-k}$ are, respectively, the low-pass and high-pass FIR filters with a subsampling factor of 2. The process is further illustrated in Figure 5.
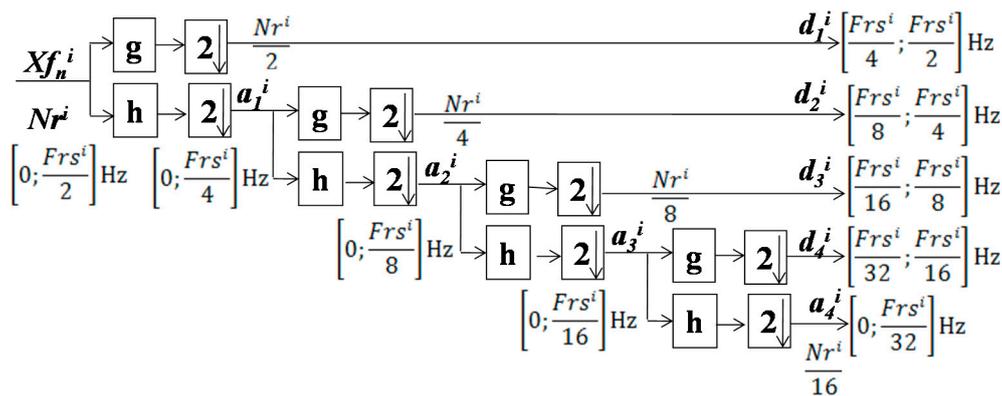


**Figure 5.** The adaptive-rate discrete wavelet transform (ARDWT) scheme with four levels of decomposition.

According to [40], sub-bands extracted by wavelet decomposition are functions of the incoming signal sampling frequency. In the proposed solution, the $W^i$ can be resampled at a specific frequency $Frs^i$, resulting in an adaptive-rate decomposition for each $W^i$, potentially achievingsub-bands with a lesser computational cost. This happens because the system has to process fewer number of samples compared to fix-rate decomposition concepts [4–6,12–15]. Furthermore, the adjustment of $Frs^i$ allows for a better focus on the incoming signal band of interest compared to the fix-rate counterparts [4–6,12–15].

$$a_m^i = \sum_{k=1}^{K_g^i} xf_n^i \cdot g_{2n-k} \tag{8}$$

$$d_m^i = \sum_{k=1}^{K_g^i} xf_n^i \cdot h_{2n-k} \tag{9}$$

### 2.6.2. Features

The wavelet coefficients obtained for each intended sub-band—$d_1^i$, $d_2^i$, $d_3^i$, $d_4^i$, and $a_4^i$—were used for extracting classifiable signal features. Nine statistical features were extracted for each considered sub-band. Therefore, in total, 45 features were employed to represent each selected segment.

The same features were extracted from each considered sub-band, as listed below.

*Power Spectrum of the Signal (PS)*

Power spectrum (PS) is computed as the average absolute value of the spectral means.

*Mean Absolute Value of the Signal (MAV)*

The mean absolute value (MAV) is calculated by adding all the absolute values of coefficients and then normalizing the sum.

*Standard Deviation (STD)*

Standard deviation (STD) is the measure of intended coefficients dispersion from the mean value.

*Skewness of the Signal (SK)*

Skewness (SK) is a measure of the asymmetry of the frequency distribution around its mean.

*Kurtosis of the Signal (K)*

Kurtosis (K) is a measure of the curvature of the considered coefficients.

*Mean Ratio (R)*

The mean ratio (R) is the ratio of the mean value of the detailed signal to the mean value of the approximate signal.

*Peak Positive Value (PV)*

The peak positive value (PV) is the maximum positive amplitude of the considered coefficients.

*Peak Negative Value (NV)*

The peak negative value (NV) is the maximum negative amplitude of the considered coefficients.

*Second Peak Negative Value (NV2)*

The second peak negative value (NV2) is the second maximum negative amplitude of the considered coefficients.

### 2.7. Classification Methods

Here, once relevant features were extracted, the data were represented in the form of a reduced data matrix, composed of 45 features, with 9 from each of the 5 selected sub-bands for each intended instance. Since the employed dataset had 5 ECG classes with 300 instances per class, the resulting data matrix had a size of 1500 × 45. To classify this data matrix, the following classification techniques were employed.

Several of the used classification techniques require parameters tuning. To do this, we used the standard method of validation to tune parameters during the training phase to get the appropriate average result. These values were then fixed for the testing phase and are mentioned below.

### 2.7.1. *k*-Nearest Neighbors (*k*-NN)

The *k*-nearest neighbors (*k*-NN) algorithm [41,42] uses the *k* nearest neighbors of a test sample from the training dataset. Let $v_j$ represent a sample and $<v_j, l_j>$ denote a tuple of a training sample and

its label, $l_j \in [1,C]$ where $C$ is the number of classes. Given a test sample, $z$, the mathematical process of computing the nearest neighbor, $j$, is presented by Equation (10):

$$\underset{j}{\operatorname{argmin}}\ dist\big(d_j, t\big)\forall j = 1..N \tag{10}$$

In the designed solution, $k = 3$ and the Euclidean distance metric was used. The final label of $z$ was selected as the most frequent label of the $k$ chosen neighbors.

### 2.7.2. Artificial Neural Network (ANN)

The artificial neural network (ANN) is a popular class of algorithms and is based on the perceived mechanics of the human brain [43]. Standard multi-layered perceptron (MLP) was used here. The input and output layers, respectively, corresponded to the 45 features and the 5 possible classes. Hidden layers are important to model complex data, but care must be taken to avoid over-fitting. The number of the hidden layers was chosen equal to 5. The training function property 'traingdx,' in MATLAB [44] was used. This set the learning to gradient descent momentum and set a variable learning rate. The activation function used was the radial basis, and the maximum epochs were set to 1000.

### 2.7.3. Support Vector Machines (SVMs)

The support vector machine (SVM) was developed by Cortes and Vapnik [45]. It searches for the optimal separating hyperplane between support vectors of two classes. The process of separating the hyperplane is described mathematically in Equation (11), where $x$ is the sample vector $x = [x_1, x_2 \ldots x_p]$ with $p$ attributes, $w = [w_1, w_2 \ldots w_p]$ is the weights vector, and $b$ is a scalar bias.

$$h(x) = sign\big(w{\cdot}x^{\mathrm{T}} + b\big) \tag{11}$$

For categorizing multiple classes, different approaches can be used [46,47], such as the one-vs-all approach or the one-vs-one approach. In this solution, the one-vs-all strategy was used with sequential minimal optimization (SMO) to train the classifier weights. The kernel function used was the polynomial of order 3, and the regularization parameter was set to 50.

### 2.7.4. Random Forest (RF)

The random forest (RF) was developed by using ideas put forward by Ho Tim Kam [48]. It is a technique that takes advantages of multiple classifiers. It constructs a multitude of decision trees at the training time and use the output from these trees to form a consensus. In contrast with bagging, RF may employ different decision tree techniques for the multiple subsets created. In the designed solution, the number of trees was set equal to 60, and out-of-bag predictions were retained for each tree. For the split at the nodes, the interaction-curvature method was selected, as it minimized the $p$-value of the chi-square tests of independence between each predictor and the response. The number of splits on each branch was limited to 10.

### 2.7.5. Bagging (BG)

Bagging (BG) is a bootstrap aggregation of classification trees. Multiple trees are allowed to fit the training data so that any bias, such as over-fitting, can be dealt with by using the ensemble of trees. Hence, bagging can be a powerful tool in classification. In this study, the number of bagged trees was set to 50, and ensemble predictions were used for out-of-bag observations.

*2.8. Performance Evaluation Metrics*

2.8.1. Compression Ratio

The compression ratio compares the proposed system performance in terms of reduction in the amount of information to be transmitted and classified compared to the conventional approach. In the classical case, acquired ECG data points are transmitted to the cloud without performing any features selection [24]. If $N_r$ and $P$ are, respectively, the count of data points to be classified in the conventional and the proposed approach, then the compression ratio, $R_{COMP}$, can be calculated by using Equation (12):

$$R_{COMP} = \frac{N_r}{P} \tag{12}$$

2.8.2. Computational Complexity

The computational complexity of the embedded processing chain till the denoising module was studied in detail. The complexity of wavelet decomposition, features extraction, and cloud processing modules was analyzed at an abstract level by considering the reduction in amount of information to be processed by these modules.

The resampled signal was denoised by using an enhanced adaptive-rate FIR (ARFIR) filtering concept [27]. The arithmetic complexity of a classical $K$ order FIR filter is clear from Equation (5). It executes $(K-1)$ additions and $K$ multiplications while calculating an output sample. Therefore, for $N$ samples, the entire computational complexity $C_{FIR}$ can be calculated by using Equation (13):

$$C_{FIR} = \underbrace{(K-1){\cdot}N}_{Additions} + \underbrace{K{\cdot}N}_{Multiplications} \tag{13}$$

For the suggested solution, the online filter selection and the selected segment resampling processes necessitated additional operations.

- Filter selection for $W^i$ wasresolved by using the successive approximation algorithm. Therefore, resolving the value of $Frs^i$ for $Q$ reference filters, in the worst case, requires $log_2(Q)$ comparisons [29].
- Resampling wasrealized by using the SLI. For $W^i$, the complexity of SLI was $Nr^i$ additions and $Nr^i$ binary weighted right shifts. The complexity of binary weighted right shift wasnegligible compared to the addition and multiplication processes [49]. Therefore, it wasignored.
- The complexity of the $K^i$ order FIR filtering for $Nr^i$ samples could be calculated as: $(K^i-1){\cdot}Nr^i$ additions and $K^i{\cdot}Nr^i$ multiplications.

Therefore, the overall complexity of the proposed ARFIR method for $W^i$ can be calculated by using Equation (14):

$$C_{ARFIR} = \underbrace{(K^i-1){\cdot}Nr^i + Nr^i + log_2(Q)}_{Additions} + \underbrace{K^i{\cdot}Nr^i}_{Multiplications} \tag{14}$$

2.8.3. Classification Precision

The proposed solution seems appealing in terms of hardware complexity, compression, processing, and transmission efficiencies. However, it can lag in terms of precision. Therefore, the performance of the whole system wasstudied in terms of its classification accuracy. To avoid any bias in estimating the classification performance due to the limited dataset, cross-validation schema has been popularly used in the literature [47]. Therefore, 10-fold cross-validation was used in this study. All tested algorithms are provided with the same dataset, both training and test, for each fold. Similarly, to avoid any biasness in findings from any one measure, the following evaluation measures were used.

Accuracy (Acc)

Accuracy (Acc) is the percentage of labels that have been correctly classified. Let TP, TN, FP, and FN, respectively, denote true positives, true negatives, false positives, and false negatives in the predicted labels. Then, the mathematical formulation for accuracy is given by Equation (15). The accuracy measure results in a value between 0 and 1, with a higher value signifying a better performance.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{15}$$

Normalized Mutual Information (NMI)

Normalized Mutual Information (NMI) is an information theoretic score calculated as the mutual information between two distributions-labels predicted by a classifier and the real labels of the data. The value of NMI scores ranges between 0 and 1, with a higher score signifying a better classification. It is mathematically computable by using Equation (16), where $X$ is the predicted clustering labels via the algorithm, $Y$ is the real labels from the data, $k(X)$ denotes the number of clusters in $X$, $r_j^X$ is the number of elements in cluster $j$ according to $X$, $r_{uj}$ is the number of elements predicted in $u$ but actually belonging to $j$, and $n$ is the total number of elements.

$$MI = \frac{\sum_1^{k(X)} \sum_1^{k(Y)} r_{uj} log_{k(X) \cdot k(Y)} \left( \frac{n \cdot r_{uj}}{r_u^X r_j^Y} \right)}{\sqrt{\left( \sum_1^{k(X)} r_u^X log \left( \frac{r_u^X}{n} \right) \right) \left( \sum_1^{k(Y)} r_u^Y log \left( \frac{r_u^Y}{n} \right) \right)}} \tag{16}$$

*F*-Measure (F1)

The *F*-measure (F1) balances the values of precision and recall. We usually talk about a macro (without taking class sized into account) or micro (taking class sizes into account) *F*-measure. However, as all classes had the same data size in the studied case, we simply employ the *F*-measure. Mathematically, the *F*-measure is expressed by Equation (17), where *precision* $= \frac{TP}{(TP+FP)}$ and *recall* $= \frac{TP}{(TP+FN)}$.

$$F = \frac{2 * precision * recall}{precision + recall} \tag{17}$$

Kappa Index (Kappa)

The kappa index (kappa) is a widely used statistics to judge the agreement of two clustering results. It is usually considered more robust than simple accuracy because it takes the possibility of agreement by chance into account. The most popular version is the Cohen's kappa measure, which is mathematically expressed by Equation (18), where $p_0$ is the percentage of agreement between the predicted and actual labels, similar to accuracy, and $p_e$ is the hypothetical probabilistic chance of such an agreement occurring randomly. It is given as $p_e = \frac{(TP+TN)(TP+FN)+(FP+TN)(FP+FN)}{(TP+TN+FP+FN)^2}$. A perfect classification gives *kappa* $= 1$, and if the classification is merely due to chance, *kappa* $= 0$.

$$kappa = 1 - \frac{1 - p_0}{1 - p_e} \tag{18}$$

Specificity (Sp)

Specificity (Sp) measures a test's ability to correctly generate a negative result for instances thatdo not belong to a given class. It is also known as the "true negative" rate. It is expressed by Equation (19):

$$Sp = \frac{TN}{TN + FP} \tag{19}$$

## 3. Results

The performance of the suggested solution wasstudied for five arrhythmia classes [36]. All system modules were implemented and validated by using the MATLAB® [44]. Examples of the pre-segmented heartbeats from all considered classes are shown in Figure 6.
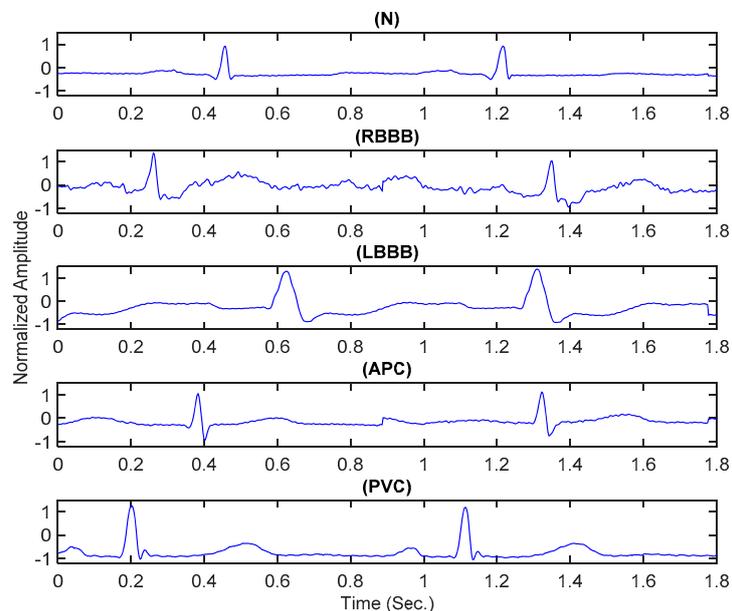
**Figure 6.** Examples of the pre-segmented electrocardiogram (ECG) signals.

In the classical case, the signal is band-limited to 60 Hz, and it is acquired with traditional ADC of 11-bit resolution at a sampling rate of 360 Hz [38]. The signal is divided into segments of 0.9 s, and each segment consists of 320 samples.

In the proposed case, to test the LCADC, the considered ECG signals were reconstructed by using $U = 400$. The reconstructed signals were acquired with an LCADC of $M = 5$-bit. The ECG signals were band-limited up to $f_{max} = 60$ Hz. Therefore, the maximum LCADC sampling frequency was $Fs_{max} = 3.6$ kHz. Examples of the pre-segmentation ECG signal, from the PVC class, acquired with an 11-bit classical ADC and a 5-bit LCADC, are shown in Figure 7.
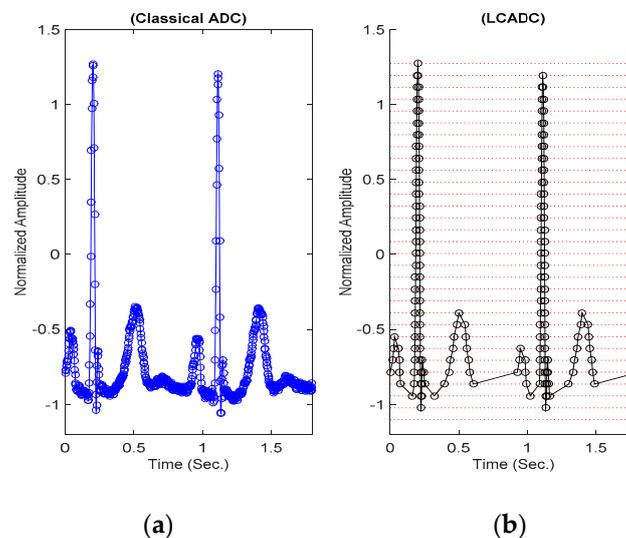
(**a**)											(**b**)

**Figure 7.** The pre-segmented ECG signal, acquired with an 11-bit resolution conventional ADC (**a**) and acquired with a 5-bit resolution LCADC (**b**).

The LCADC output was segmented by using the ASA. The ASA adapted $L^i$ and $Frs^i$ according to the $\widetilde{x}(t)$ temporal variations. It contributed to the enhancement of the system's computational efficiency. The average compression ratios were computed for all 300 instances of each considered class. These were, respectively, 3.13-, 2.86-, 3.01-, 3.02-, and 3.05-fold for classes N, RBBB, LBBB, APC, and PVC. The attained overall average reduction in the number of collected samples for all five classes was 3.01-fold.

The resampled signal wasdenoised by using the ARFIR filtering technique [29]. A band-pass filter bank was designed for the cut-off frequencies of $[F_{Cmin} = 0.7; F_{Cmax} = 35]$ Hz. The Filter bank was implemented for a set of sampling frequencies, $Fref$, between $Fs_{min} = 75$ Hz $> 2 \cdot F_{Cmax}$ to $F_r = 360$ Hz. In this case, $\Delta = 19$ Hz was chosen. It realized a bank of $Q = 16$ band-pass filters. A summary of the designed filter bank is shown in Table 2.

**Table 2.** Summary of the reference filters bank parameters.

| $hc_k$ | $h1_k$ | $h2_k$ | $h3_k$ | $h4_k$ | $h5_k$ | $h6_k$ | $h7_k$ | $h8_k$ | $h9_k$ | $h10_k$ | $h11_k$ | $h12_k$ | $h13_k$ | $h14_k$ | $h15_k$ | $h16_k$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $Fref_c$ **(Hz)** | 75 | 94 | 113 | 132 | 151 | 170 | 189 | 208 | 227 | 246 | 265 | 284 | 303 | 322 | 341 | 360 |
| $K_c$ | 23 | 30 | 36 | 43 | 49 | 55 | 61 | 68 | 74 | 80 | 86 | 92 | 99 | 105 | 111 | 117 |

Denoising enhanced the intended signal SNR and resulted in a better classification [25]. The online filter order adaption resulted in signal denoising with a reduced computational cost compared to the time-invariant traditional ones [29]. The gains of the proposed ARFIR were computed over the classical one by using Equations (13) and (14). A summary of results for all 300 instances for each considered ECG class is presented in Tables 3 and 4.

**Table 3.** Summary of the adaptive-rate finite impulse response (ARFIR) computational gains, in additions, over the classical FIR filtering.

| CLASS | Max Gain in Additions | Min Gain in Additions | Mean Gain in Additions | Median Gain in Additions |
|---|---|---|---|---|
| Normal | 15.97 | 6.67 | 8.12 | 7.90 |
| RBBB | 12.17 | 6.66 | 7.32 | 7.27 |
| LBBB | 12.46 | 6.70 | 7.86 | 7.66 |
| APC | 10.75 | 6.75 | 7.71 | 7.56 |
| PVC | 15.06 | 6.71 | 8.03 | 7.77 |

**Table 4.** Summary of the ARFIR computational gains, in multiplications, over the classical FIR filtering.

| CLASS | Max Gain in Multiplications | Min Gain in Multiplications | Mean Gain in Multiplications | Median Gain in Multiplications |
|---|---|---|---|---|
| Normal | 16.14 | 6.74 | 8.31 | 8.09 |
| RBBB | 12.30 | 6.73 | 7.48 | 7.43 |
| LBBB | 12.58 | 6.76 | 8.04 | 7.84 |
| APC | 10.86 | 6.81 | 7.89 | 7.74 |
| PVC | 15.22 | 6.76 | 8.22 | 7.95 |

Tables 3 and 4 illustrate that the proposed ARFIR technique achieved a noticeable computational gain over the traditional counterpart. It was attained by intelligently adapting the system parameters like $Frs^i$, $Nr^i$, and $hc_k$ for each selected segment ($W^i$) by following the $\widetilde{x}(t)$ temporal variations.

Each classifier was adjusted during the training phase with parameters set as mentioned in Section 2.7. Five classifiers, namely ANN, $k$-NN, SVM, RF, and bagging, were used. Their parameters were adjusted according to methods described in Section 2.7. They classified the intended dataset, and we measured the different evaluation metrics. The results are summarized in Table 5.

A summary of detail classifications results, in terms of TP, FP, FN, and TN for the five considered classes, is presented in Figure 8. These results showed that the integration of adaptive-rate signal

acquisition and processing chain with RF and bagging resulted in superior classification performances as compared to *k*-NN, ANN, and SVM. Since RF uses multiple classifiers, it is less likely to be biased. This was reflected in the evaluation where RF resulted in the best NMI score and kappa statistics. On the other hand, *k*-NN can be easily biased by the chosen neighbors, particularly if there are outliers in the data. This was reflected in its low score across all indices. Overall, the highest accuracy score of 0.97 was attained by the designed solution.

**Table 5.** Performance of different classifiers using five evaluation metrics on the test dataset. ANN: artificial neural network; *k*-NN: *k*-nearest neighbors; SVM: support vector machine; and RF: random forest.

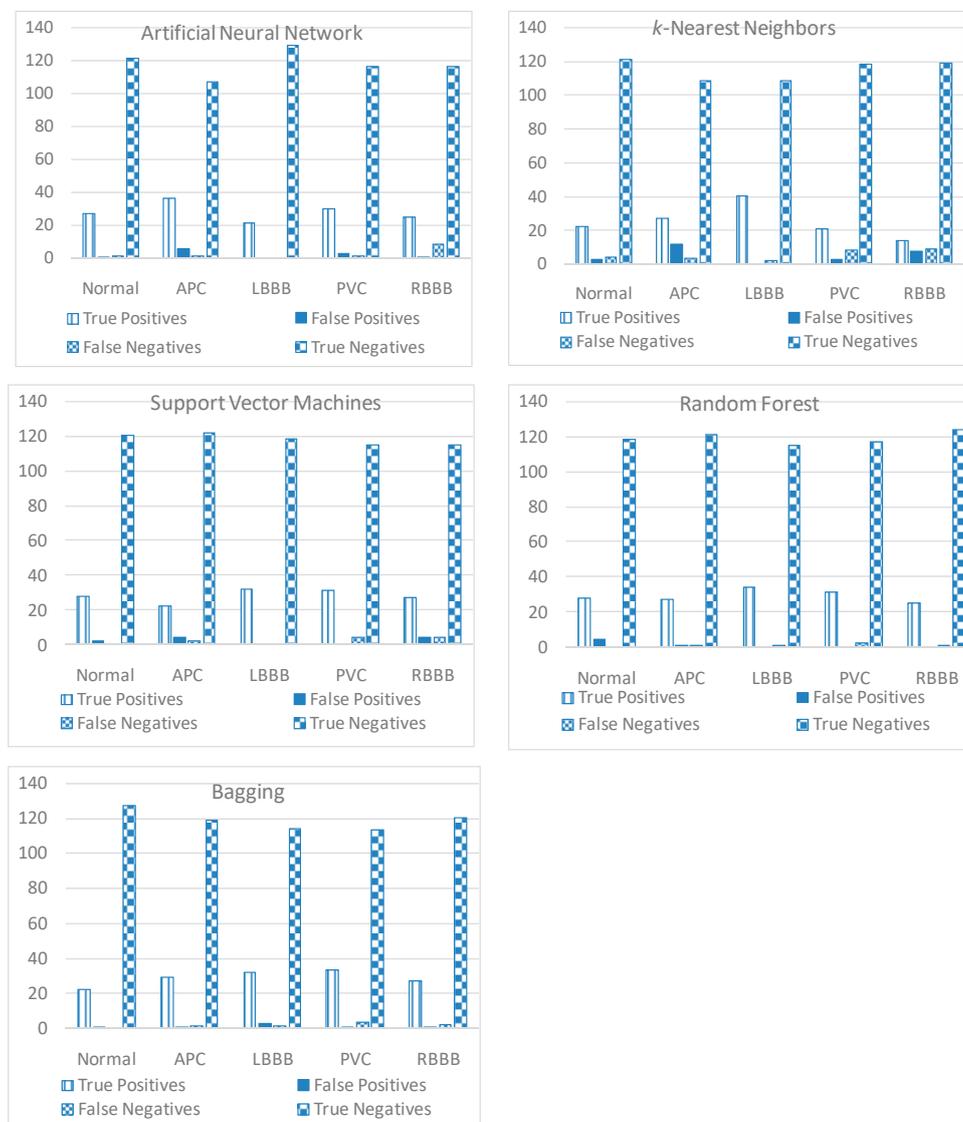| Classifier | Acc | NMI | F1 | Kappa | Sp |
|---|---|---|---|---|---|
| ANN | 0.93 | 0.82 | 0.93 | 0.77 | 0.98 |
| *k*-NN | 0.83 | 0.71 | 0.81 | 0.46 | 0.96 |
| SVM | 0.93 | 0.84 | 0.93 | 0.75 | 0.98 |
| RF | 0.97 | 0.94 | 0.97 | 0.90 | 0.99 |
| Bagging | 0.96 | 0.89 | 0.96 | 0.79 | 0.99 |



**Figure 8.** True positive (TP), false positive (FP), false negative (FN), and true negative (TN) for different ECG classes.

For further analysis, we considered the detail classification results for each intended class, as shown in Figure 8. The overall false positive and false negative counts by using RF were the lowest compared to bagging, k-NN, ANN, and SVM, thus confirming the outperformance of RF compared to other considered algorithms. The most difficult class to discriminate was the RBBB, which was confused with LBBB, while the easiest class to discriminate was normal compared to the other four classes.

## 4. Discussion

The appealing features of the suggested solution are clear from results, presented in Section 3. This solution was attained by intelligently exploiting level-crossing sampling, adaptive-rate processing, and robust classifiers. The values of $Nr^i$ were adapted by the ASA. We showed how $Fs^i$, $L^i$, and $Frs^i$ were adjusted as a function of the temporal variations of $\widetilde{x}(t)$. The adjustment of $Frs^i$ avoided unnecessary interpolations during the resampling process [29]. This signal-driven, adaptive-rate sampling resulted in a three-fold decrease in the number of collected samples compared to classical counterparts. It guaranteed a noticeable reduction in the processing load of the post denoising and features extraction.

$K^i$ represents how the adjustment of $hc_k$ for $W^i$ avoids the unnecessary operations while conditioning the selected segments. It yielded a noteworthy computational gain of the proposed denoising method over the conventional counterparts. The average gains in additions and multiplications, for the considered 1500 instants, of the employed ARFIR over the conventional one were 7.81- and 7.99-fold, respectively. Additionally, the adaptation of $Frs^i$ resulted in an adaptive-rate sub-band decomposition. $Frs^i$ reflected the $\widetilde{x}(t)$ frequency content [35]. Therefore, it was able to result in an improved focus on the band of interest, as the decomposition process is a function of $Frs^i$ [40]. The decomposition was attained by using the half-band FIR filters. A three-fold decrease in the number of collected samples confirmed computationally efficient sub-band decomposition compared to the counter fix-rate solutions [4,6,12,13,29,40].

Originally, each intended instance was composed of 320 samples. After features extraction, their dimensionality was reduced to 45 features, resulting in a 7.1-fold dimensionality reduction and guaranteeing the same factor of reduction in the data transmission activity, bandwidth, and power consumption. Additionally, on the cloud side, the processing of 7.1-fold less amount of information assured a similar factor of processing and resource utilization gain during the classification process.

The adoption of performing most of the signal processing tasks via the front-end processor and transmitting only the extracted features to the cloud was also beneficial in realizing an optimized front-end ECG wearable device while keeping the system configurable. Moreover, in the designed framework, the signal was digitized with a $M = 5$-bit resolution LCADC. In the classical case, the digitization is performed with an $M = 11$-bit resolution ADC [4–6,12–15]. Our study confirmed a simpler circuit level realization compared to the traditions ones. A low resolution ADC can also be used in the classical case, but it can reduce the system SNR [38] and can therefore degrade classification performance. On other hand, the effective resolution of LCADC is fairly independent of $M$ [38]. Therefore, an accuracy of 97% was achieved by the proposed solution, even when using a 5-bit resolution LCADC.

The idea of embedding the level-crossing sampling and adaptive-rate processing in ECG signal-based automatic arrhythmia diagnosis is a novel concept. In [16,24,30–32], new approaches were proposed for effective ECG acquisition. These mainly focused on the analog-to-digital conversion of ECG signals. In [30,31], the authors demonstrated that how the use of LCS-based architectures led towards a simplified and power efficient analog/digital (A/D) conversion. In [32], the authors showed how the proposed ADC could adjust its effective resolution as a function of the activity of the input signal. In [16,24], efficient and real-time QRS detection mechanisms were proposed.

In contrast to [16,24,30–32], the proposed solution did not only focus on the design of an effective LCADC for ECG signal acquisition or QRS complexes detection; it additionally presented an application of the ASA and the adaptive-rate resampling, denoising, and sub-band decomposition approaches

on the LCADC output. It was performed to achieve efficient segmentation, noise reduction, and features extraction of digitized ECG pulses to well prepare them for post cloud-based automatic arrhythmia classification.

The proposed technique is original, and comparing it with existing state-of-the-art methods was not straight forward, since they are based on classical sampling and processing. Additionally, each study uses a different number of subjects and different types of arrhythmia classes. It is also delicate to compare the diversity of classification methods and techniques for ECG signal processing. However, a comparison was performed with previous studies that use the same ECG dataset and DWT-based features extraction. The highest accuracies of classification for all considered studies are presented in Table 6, which shows that the proposed solution attained an analogous or better classification accuracy as compared to the fix-rate counterparts [4–6,12–15] while assuring noticeable processing, transmission, and hardware simplicity gains.

**Table 6.** Comparison with state-of-the-art methods.

| Study | Features Extraction | Classification Method | No. of Classes | Accuracy (%) |
|---|---|---|---|---|
| [4] | Wavelet Packet Decomposition (WPD) and Wavelet-Based kernel Principle Component Analysis (wkPCA) | Backpropagation Neural Network (BNN) | 5 | 98.03 |
| [5] | Wavelet Packet Entropy (WPE) | Random Forests (RF) | 5 | 94.61 |
| [6] | Discrete Wavelet Transform (DWT) | Probabilistic Neural Network (PNN) | 8 | 92.75 |
| [12] | Discrete Wavelet Transform (DWT), Temporal, and Morphological | Support Vector Machine (SVM) | 4 | 98.39 |
| [13] | Discrete Wavelet Transform (DWT) and Principle Component Analysis (PCA) | Support Vector Machine (SVM)-Radial Basis Function (RBF) | 5 | 96.92 |
| [14] | Bispectrumand Principle Component Analysis (PCA) | Support Vector Machine (SVM)-Radial Basis Function (RBF) | 5 | 93.48 |
| [15] | HermiteFunction Coefficient and Temporal Features | Optimized block-based Neural Network (OBNN) | 5 | 97.00 |
| This Study | Wavelet Decomposition and Sub-band statistical features | Random Forests (RF) | 5 | 97.00 |

The main advantage of the proposed solution over the previous ones is the elimination of unnecessary samples to process and to introduce a real-time compression gain in the system. This was achieved by tactfully embedding the level-crossing sampling and ASA into the system. Similar gains could be achieved by embedding these concepts in counterparts [4–6,12–15].

## 5. Conclusions

An original level-crossing ECG signal sampling, adaptive-rate denoising, and sub-band decomposition and features extraction approach was proposed. It was shown that the proposed framework achieved a three-fold reduction in the number of acquired samples, thus leading to a significant reduction in the computational complexity of the designed system compared to the conventional counterparts. The overall average gains in additions and multiplications for the designed adaptive-rate denoising module were computed as 7.81- and 7.99-fold, respectively. The adaptive-rate processing promised a similar factor of processing gain during the adaptive-rate sub-band decomposition. The proposed features extraction concept reduced the incoming data dimensionality with a factor of 7.1 and assured a remarkable gain in terms of the power consumption and transmission activity between the ECG wearable device and cloud-based classification. Moreover, a similar magnitude of processing gain was also evident in the post cloud-based classifier because it had to deal with a 7.1 times lesser amount of information. The highest classification accuracy of 97% was attained by the random forest, which is comparable, and in some cases better, to existing state-of-the-art solutions. The plus side is our gain in compression, processing, and transmission

bandwidth, which show the potential of using the suggested solution for the design and development of low power and efficient ECG wearable devices in the mobile healthcare framework.

## References

1. Brenyo, A.; Aktas, M.K. Review of complementary and alternative medical treatment of arrhythmias. *Am. J. Cardiol.* **2014**, *113*, 897–903. [CrossRef] [PubMed]

2. Banerjee, S.; Mitra, M. Application of cross wavelet transform for ECG pattern analysis and classification. *IEEE Trans. Instrum. Meas.* **2013**, *63*, 326–333. [CrossRef]

3. Mozaffarian, D.; Benjamin, E.J.; Go, A.S.; Arnett, D.K.; Blaha, M.J.; Cushman, M.; Howard, V.J. Heart disease and stroke statistics-2016 update a report from the American Heart Association. *Circulation* **2016**, *133*, e38–e48. [CrossRef] [PubMed]

4. Hai, N.T.; Nguyen, N.T.; Nguyen, M.H.; Livatino, S. Wavelet-based kernel construction for heart disease classification. *Adv. Electr. Electr. Eng.* **2019**, *17*, 306–319. [CrossRef]

5. Li, T.; Zhou, M. ECG classification using wavelet packet entropy and random forests. *Entropy* **2016**, *18*, 285. [CrossRef]

6. Gutiérrez-Gnecchi, J.A.; Morfin-Magana, R.; Lorias-Espinoza, D.; del Carmen Tellez-Anguiano, A.; Reyes-Archundia, E.; Méndez-Patiño, A.; Castañeda-Miranda, R. DSP-based arrhythmia classification using wavelet transform and probabilistic neural network. *Biomed. Signal Process. Control* **2017**, *32*, 44–56. [CrossRef]

7. Sharma, R.R.; Pachori, R.B. Baseline wander and power line interference removal from ECG signals using eigenvalue decomposition. *Biomed. Signal Process. Control* **2018**, *45*, 33–49. [CrossRef]

8. Hesar, H.D.; Mohebbi, M. ECG denoising using marginalized particle extended kalman filter with an automatic particle weighting strategy. *IEEE J. Biomed. Health Inform.* **2016**, *21*, 635–644. [CrossRef]

9. Singhal, A.; Singh, P.; Fatimah, B.; Pachori, R.B. An efficient removal of power-line interference and baseline wander from ECG signals by employing Fourier decomposition technique. *Biomed. Signal Process. Control* **2020**, *57*, 101741. [CrossRef]

10. Sahoo, S.; Kanungo, B.; Behera, S.; Sabut, S. Multiresolution wavelet transform based feature extraction and ECG classification to detect cardiac abnormalities. *Measurement* **2017**, *108*, 55–66. [CrossRef]

11. Martis, R.J.; Acharya, U.R.; Min, L.C. ECG beat classification using PCA, LDA, ICA and discrete wavelet transform. *Biomed. Signal Process. Control* **2013**, *8*, 437–448. [CrossRef]

12. Martis, R.J.; Acharya, U.R.; Mandana, K.; Ray, A.; Chakraborty, C. Cardiac decision making using higher order spectra. *Biomed. Signal Process. Control* **2013**, *8*, 193–203. [CrossRef]

13. Shadmand, S.; Mashouf, B. A new personalized ECG signal classification algorithm using block-based neural network and particle swarm optimization. *Biomed. Signal Process. Control* **2016**, *25*, 12–23. [CrossRef]

14. Zhang, X.; Lian, Y. A 300-mV 220-nW event-driven ADC with real-time QRS detection for wearable ECG sensors. *IEEE Trans. Biomed. Circ. Syst.* **2014**, *8*, 834–843. [CrossRef]

15. de Ruvo, E.; Sciarra, L.; Martino, A.M.; Rebecchi, M.; Iulianella, R.V.; Sebastiani, F.; Fagagnini, A.; Borrelli, A.; Scarà, A.; Grieco, D.; et al. A prospective comparison of remote monitoring systems in implantable cardiac defibrillators: Potential effects of frequency of transmissions. *J. Interv. Card. Electrophysiol.* **2016**, *45*, 81–90. [CrossRef]

16. Guzik, P.; Malik, M. ECG by mobile technologies. *J. Electrocardiol.* **2016**, *49*, 894–901. [CrossRef]

17. Miah, S.J.; Gammack, J.; Hasan, N. Extending the framework for mobile health information systems Research: A content analysis. *Inf. Syst.* **2017**, *69*, 1–24. [CrossRef]

18. Rezaii, T.Y.; Beheshti, S.; Shamsi, M.; Eftekharifar, S. ECG signal compression and denoising via optimum sparsity order selection in compressed sensing framework. *Biomed. Signal Process. Control* **2018**, *41*, 161–171. [CrossRef]

19. Singh, A.; Sharma, L.N.; Dandapat, S. Multi-channel ECG data compression using compressed sensing in eigenspace. *Comput. Biol. Med.* **2016**, *73*, 24–37. [CrossRef]

20. Chen, S.; Hua, W.; Li, Z.; Li, J.; Gao, X. Heartbeat classification using projected and dynamic features of ECG signal. *Biomed. Signal Process. Control* **2017**, *31*, 165–173. [CrossRef]

21. Niederhauser, T.; Haeberlin, A.; Jesacher, B.; Fischer, A.; Tanner, H. Model-based delineation of non-uniformly sampled ECG signals. In Proceedings of the 2017 Computing in Cardiology (CinC), Rennes, France, 24–27 September 2017; pp. 1–4. [CrossRef]

22. Tekeste, T.; Saleh, H.; Mohammad, B.; Ismail, M. Ultra-Low Power QRS Detection and ECG Compression Architecture for IoT Healthcare Devices. *IEEE Trans. Circ. Syst. Regul. Pap.* **2018**, *66*, 669–679. [CrossRef]

23. Qaisar, S.M.; Subasi, A. An adaptive rate ECG acquisition and analysis for efficient diagnosis of the cardiovascular diseases. In Proceedings of the 2018 IEEE 3rd International Conference on Signal and Image Processing (ICSIP), Shenzhen, China, 13–15 July 2018; pp. 177–181. [CrossRef]

24. Budiman, E.S. Multi-Rate Analyte Sensor Data Collection with Sample Rate Configurable Signal Processing. U.S. Patent 20170027483A1, 2017.

25. Qaisar, S.M.; Fesquet, L.; Renaudin, M. Adaptive rate filtering a computationally efficient signal processing approach. *Signal Process.* **2014**, *94*, 620–630. [CrossRef]

26. Qaisar, S.M.; Ben-Romdhane, M.; Anwar, O.; Tlili, M.; Maalej, A.; Rivet, F.; Rebai, C.; Dallet, D. Time-domain characterization of a wireless ECG system event driven A/D converter. In Proceedings of the 2017 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), Torino, Italy, 22–25 May 2017; pp. 1–6. [CrossRef]

27. Qaisar, S.M. Efficient mobile systems based on adaptive rate signal processing. *Comput. Electr. Eng.* **2019**, *79*, 106462. [CrossRef]

28. Hou, Y.; Qu, J.; Tian, Z.; Atef, M.; Yousef, K.; Lian, Y.; Wang, G. A 61-nW level-crossing ADC with adaptive sampling for biomedical applications. *IEEE Trans. Circ. Syst. II Express Briefs* **2018**, *66*, 56–60. [CrossRef]

29. Marisa, T.; Niederhauser, T.; Haeberlin, A.; Wildhaber, R.A.; Vogel, R.; Goette, J.; Jacomet, M. Pseudo asynchronous level crossing ADC for ECG signal acquisition. *IEEE Trans. Biomed. Circuits Syst.* **2017**, *11*, 267–278. [CrossRef] [PubMed]

30. Nasserian, M.; Peiravi, A.; Moradi, F. An adaptive-resolution signal-specific ADC for sensor-interface applications. *Analog Integr. Circ. Signal Process.* **2019**, *98*, 125–135. [CrossRef]

31. Miskowicz, M. *Event-Based Control and signal Processing*; CRC Press: Boca Raton, FL, USA, 2015. [CrossRef]

32. Mashhadi, M.B.; Salarieh, N.; Farahani, E.S.; Marvasti, F. Level crossing speech sampling and its sparsity promoting reconstruction using an iterative method with adaptive thresholding. *IET Signal Process.* **2017**, *11*, 721–726. [CrossRef]

33. Moser, B.A.; Lunglmayr, M. On quasi-isometry of threshold-based sampling. *IEEE Trans. Signal Process.* **2019**, *67*, 3832–3841. [CrossRef]

34. Moody, G.B.; Mark, R.G. The impact of the MIT-BIH arrhythmia database. *IEEE Eng. Med. Biol. Mag.* **2001**, *20*, 45–50. [CrossRef]

35. Qaisar, S.M.; Yahiaoui, R.; Gharbi, T. An efficient signal acquisition with an adaptive rate A/D conversion. In Proceedings of the 2013 IEEE International Conference on Circuits and Systems (ICCAS), Kuala lumpur, Malaysia, 18–19 September 2013; pp. 124–129. [CrossRef]

36. Qaisar, S.M.; Akbar, M.; Beyrouthy, T.; Al-Habib, W.; Asmatulah, M. An error measurement for resampled level crossing signal. In Proceedings of the 2016 Second International Conference on Event-based Control, Communication, and Signal Processing (EBCCSP), Krakow, Poland, 13–15 June 2016; pp. 1–4. [CrossRef]

37. Massopust, P. *Interpolation and Approximation with Splines and Fractals*; Oxford University Press, Inc.: New York, NY, USA, 2010; ISBN 978-0-19-533654-2.

38. Zhang, Z.; Telesford, Q.K.; Giusti, C.; Lim, K.O.; Bassett, D.S. Choosing wavelet methods, filters, and lengths for functional brain network construction. *PLoS ONE* **2016**, *11*, e0157243. [CrossRef] [PubMed]

39. Hussain, S.F.; Bisson, G. Text categorization using word similarities based on higher order co-occurrences. In Proceedings of the 2010 SIAM International Conference on Data Mining (SDM), Columbus, OH, USA, 29 April–1 May 2010; pp. 1–12. [CrossRef]

40.  Dudani, S.A. The distance-weighted k-nearest-neighbor rule. *IEEE Trans. Syst. Man Cybern.* **1976**, 325–327. [CrossRef]

41.  Parisi, G.I.; Kemker, R.; Part, J.L.; Kanan, C.; Wermter, S. Continual lifelong learning with neural networks: A review. *Neural Netw.* **2019**, *113*. [CrossRef] [PubMed]

42.  Guo, H.; Wang, W. Granular support vector machine: A review. *Artif. Intell. Rev.* **2019**, *51*, 19–32. [CrossRef]

43.  Hussain, S.F.; Suryani, A. On retrieving intelligently plagiarized documents using semantic similarity. *Eng. Appl. Artif. Intell.* **2015**, *45*, 246–258. [CrossRef]

44.  Moore, H. *MATLAB for Engineers*; Pearson: London, UK, 2017; ISBN 978-0-13-458964-0.

45.  Hussain, S.F. A novel robust kernel for classifying high-dimensional data using Support Vector Machines. *Expert Syst. Appl.* **2019**, *131*, 116–131. [CrossRef]

46.  Rokach, L. Decision forest: Twenty years of research. *Inf. Fusion* **2016**, *27*, 111–125. [CrossRef]

47.  Cavanagh, J. *Computer Arithmetic and Verilog HDL Fundamentals*; CRC Press: Boca Raton, FL, USA, 2017. [CrossRef]

48.  de Chazal, P.; O'Dwyer, M.; Reilly, R.B. Automatic classification of heartbeats using ECG morphology and heartbeat interval features. *IEEE Trans. Biomed. Eng.* **2004**, *51*, 1196–1206. [CrossRef]

49.  Christov, I.; Gómez-Herrero, V.; Krasteva, G.; Jekova, I.; Gotchev, A.; Egiazarian, K. Comparative study of morphological and time-frequency ECG descriptors for heartbeat classification. *Med. Eng. Phys.* **2006**, *28*, 876–887. [CrossRef]