Arrhythmia classification from single-lead ECG signals using the inter-patient paradigm

Article *in* Computer Methods and Programs in Biomedicine · January 2021

DOI: 10.1016/j.cmpb.2021.105948

4 authors, including:

Felipe Meneguitti Dias
Instituto do Coração
10 PUBLICATIONS **23** CITATIONS
SEE PROFILE

Eduardo José da Silva Luz
Universidade Federal de Ouro Preto
46 PUBLICATIONS **1,028** CITATIONS
SEE PROFILE

Henrique Monteiro
Universidade Federal de Lavras (UFLA)
15 PUBLICATIONS **49** CITATIONS
SEE PROFILE

Arrhythmia classification from single-lead ECG signals using the inter-patient paradigm

Felipe Meneguitti Dias[a], Henrique L. M. Monteiro[a], Thales Wulfert Cabral[b], Rayen Naji[c], Michael Kuehni[d], Eduardo José da S. Luz[e]

[a]*Electrical Engineering Department, Universidade Federal de Juiz de Fora, Juiz de Fora, MG, Brazil*
[b]*Electrical Engineering Department, Universidade Estadual de Campinas, Campinas, SP, Brazil*
[c]*Medical School, Universidade Federal de Juiz de Fora, Juiz de Fora, MG, Brazil*
[d]*Illinois Institute of Technology, Chicago, IL, The United States*
[e]*Computing Department, Universidade Federal de Ouro Preto, Ouro Preto, MG, Brazil*

Abstract

Background and objectives: Arrhythmia is a heart disease characterized by the change in the regularity of the heartbeat. Since this disorder can occur sporadically, Holter devices are used for continuous long-term monitoring of the subject's electrocardiogram (ECG). In this process, a large volume of data is generated. Consequently, the use of an automated system for detecting arrhythmias is highly desirable. In this work, an automated system for classifying arrhythmias using single-lead ECG signals is proposed.
Methods: The proposed system uses a combination of three groups of features: RR intervals, signal morphology, and higher-order statistics. To validate the method, the MIT-BIH database was employed using the inter-patient paradigm. Besides, the robustness of the system against segmentation errors was tested by adding jitter to the R-wave positions given by the MIT-BIH database. Additionally, each group of features had its robustness against segmentation error tested as well.
Results: The experimental results of the proposed classification system with jitter show that the sensitivities for the classes N, S, and V are 93.7, 89.7, and 87.9, respectively. Also, the corresponding positive predictive values are 99.2, 36.8, and 93.9, respectively.
Conclusions: The proposed method was able to outperform several state-of-the-art methods, even though the R-wave position was synthetically corrupted by added jitter. The obtained results show that our approach can be employed in real scenarios where segmentation errors and the inter-patient paradigm are present.

Keywords: Electrocardiogram, ECG classification, Machine learning, Inter-patient, Segmentation error, Jitter

1. Introduction

The electrocardiogram (ECG) is a technology developed at the beginning of the twentieth century that is capable of recording the electrical activity of the heart [1]. If the ECG exhibits adverse behavior, there is a sign of a cardiac problem that must be further investigated by a trained health professional.

With the rise in the number of heart diseases, especially in poor and developing countries, the advancement of technologies for the early diagnosis of these diseases is becoming increasingly important [2, 3]. Among these complications, one of the most common heart diseases is the cardiac arrhythmia, whose characterization is established by the occurrence of irregular heartbeat. The identification of these anomalies is usually performed by using Holter equipment [4]. These devices continuously monitor the electrocardiogram, generating a considerable amount of data. Since arrhythmias can manifest intermittently, the designated diagnostic specialist should analyze an excessive volume of beats from a given patient. Thus, there is the possibility of incurring in misleading diagnoses arising from human error. Therefore, there is a need to search for solutions that enable the mitigation of these mistakes. Automating the classification process of such signals is one of the possible approaches.

Email address: felipe.dias@engenharia.ufjf.br (Felipe Meneguitti Dias)

In the literature, several works have proposed automatic ECG classification systems. In [5] and [6], for example, the authors attempted to classify ECG signals using the intra-patient paradigm by employing neural networks and support vector machines, respectively. The intra-patient paradigm, however, does not simulate a real-world scenario. In such a paradigm, in order to classify a subject's heartbeat, one would need labeled heartbeats of this person in advance [7]. To overcome this limitation, Chazal et al. [8] proposed the inter-patient paradigm. In this case, a separate set of subjects is used for building the classification system, and another is used for testing.

Since then, most of the relevant works in the literature adopted the inter-patient paradigm. In [9], the authors extracted 6 types of features for ECG classification: RR intervals, morphological, statistical, high order statistics, wavelet transform, and wavelet packet entropy. Then a hierarchical XGBoost classifier was employed for classification. In [10], the ECG signal is decomposed using a dictionary formed by Stockwell, sine, and cosine functions. Furthermore, five features are extracted from each decomposed ECG: kurtosis, standard deviation, RR interval, energy, and permutation entropy. For classification, a least-square twin support vector machine is employed. The parameters for this classifier were optimized using the artificial bee colony technique. In [11], RR intervals and morphological features were employed for ECG classification using a support vector machine classifier. In addition, the authors included a disease-specific feature selection step to improve classification accuracy. In [12], the authors extracted RR intervals and morphological features from filtered ECG signals. For classification, a linear discriminant classifier was employed to classify each heartbeat.

Nevertheless, most of these works do not consider R-wave detection errors. Ye et al. [13], however, is an exception. In Ye et al. [13], the authors accounted for errors with a standard deviation of 13.9 ms. Even so, works that evaluate QRS detection algorithms consider errors up to 50 ms or 100 ms [14, 15]. Therefore, in a real-world scenario, an automatic classification system should also account for realist errors generated by R-wave detection algorithms, such as Pan-Tompkins [16].

In this work, we propose an automatic arrhythmia classification system using single-lead ECG signals. The experiments were validated in the MIT-BIH Arrhythmia database [17] using the inter-patient paradigm proposed by [8]. Three groups of features were selected for the classification system: RR intervals, morphological values, and high order statistics. These groups of features sum up to 10 features in total. Compared to other works in the literature [12, 9] that employed advanced techniques, e.g., wavelet transform, and generated several features, the proposed feature combination is simpler and not computational expensive to extract.

Then, a linear discriminant classifier is employed due to its simplicity and robustness against class imbalance. Since this classifier only compares statistics from training data and the heartbeat under evaluation, it has low computational complexity and allows for real-time implementations. Also, our proposed method is tested against segmentation errors by adding artificially generated jitter to the R-wave positions given by the MIT-BIH database. We evaluate the robustness of our proposed system for different values of errors on the R-wave detection. Besides, we individually analyze the effect of such errors on each group of features (RR intervals, morphological, and high order statistics).

The proposed methodology is applied in a three-class configuration: Normal (N), Supraventricular ectopic beat (S), and Ventricular ectopic beat (V). Considering a uniformly distributed jitter of 18 samples, we achieved sensitivities of 93.7, 89.7, and 87.9 for the classes N, S, and V, respectively. Besides, for the positive predictive values, 99.2, 36.8, and 93.9 were obtained, respectively. Contrasting to that, without jitter, we were able to achieve 94.5, 92.5, and 88.6 for the sensitives, and 99.4, 39.9, and 94.6 for the positive predictive value for the classes N, S, and V, respectively.

The remainder of this paper is structured as follows. Section 2 shows a brief literature review on the necessary steps for ECG classification systems. In Section 3, the MIT-BIH database and the inter-patient paradigm are presented. The proposed methodology for ECG classification is described in Section 4. Furthermore, in Section 5, the experimental design and results are shown. In Section 6, a discussion on the acquired results and a comparison with other works in the literature are presented. Finally, the concluding remarks of the paper are reported in Section 7.

2. Related Work

An arrhythmia identification system has four important steps: preprocessing, heartbeat segmentation, feature extraction, and classification. For each one of these steps, several approaches have been proposed in the literature. Some of these approaches are briefly reviewed in this section.

The ECG signal is usually affected by different types of noise, e.g., baseline wander and power line interference [18]. In order to better extract useful information

from the ECG, some filtering techniques are usually employed. In [19], the authors used an extended Kalman filter. In [8], the authors used two median filters to remove the baseline wander. Then, a 12-tap low pass filter with the 3 dB point located at 35 Hz was applied to remove power line interference and high frequency noise. In [20], two FIR filters were employed. The first filter was a 12-tap low pass filter with the 3 dB point at 35 Hz. The second was a high pass filter with its 3 dB point at 1 Hz. In [21, 22, 23], wavelet transforms were used to remove ECG noise. In [24], the empirical mode decomposition was used to remove high-frequency noise and baseline wander. In [25], the adaptive Fourier decomposition method was applied to denoise ECG signals.

After filtering, the next step is to detect the ECG R-waves and then segment each heartbeat. However, the majority of the works in the literature bypass the detection step and use the R-wave position given by a labeled database [7]. An important exception is the work done in [13]. In this work, the authors artificially added jitter to the R-wave position. They modeled the jitter as a gaussian with zero mean and a standard deviation of five samples, i.e., 13.9 ms. Therefore, the authors were able to show that their methodology was robust even when the R-wave detection algorithm deviates from the true R-wave position.

Afterward, it is necessary to extract features that can distinguish different heartbeat classes. To this end, numerous features have been employed in the literature. In [8], the authors combined rhythmic features, wave duration, and morphological features. In [12], normalized RR intervals were combined with morphological features. In [26, 27], independent component analysis features were employed. In [28], the authors used high order statistics of wavelet packet decomposition coefficients for ECG classification. In [10], the authors decomposed the ECG signals using a composition of sparse dictionaries. For each component, the energy, RR intervals, standard deviation, kurtosis, and the permutation entropy were extracted.

To classify each heartbeat using the extracted features, different algorithms have been used in the literature. In [8], the authors used linear discriminant classifier due to its simplicity. In [29], the authors employed a support vector machine, and they used different weights for each class to deal with class imbalance. In [30], the authors proposed the use of the optimum-path forest classifier. The authors of [9] employed a hierarchical classification process combining weighted extreme gradient boosting classifiers and a threshold classifier.

3. Materials

This study used the MIT-BIH Arrhythmia Database [17] to evaluate the proposed methodology. This database contains 48 30-minute recordings sampled at 360 Hz obtained from 47 different subjects. Moreover, the signals were bandpass filtered at 0.1-100 Hz. Besides, each recording has two ECG derivations, but only the modified limb lead II was used in this work.

Two cardiology experts analyzed the heartbeats from the MIT-BIH database. Then, upon consensus between these two experts, a label was given for each beat. Therefore, the database contains the R-wave position for each beat and its corresponding class. Following the Association for the Advancement of Medical Instrumentation (AAMI) recommendations, four subjects with paced beats were excluded. Besides, the 18 types of heartbeats were clustered in five groups: Normal (N), Supraventricular ectopic beat (S), Ventricular ectopic beat (V), Fusion beat (F) and Unknown beat (Q). Table 1 shows the mapping between MIT-BIH and AAMI labels.

Due to the small number of events in classes F and Q, both of them were excluded from this work. Besides, following the recommendation given by [8], the 44 recordings of the MIT-BIH database were divided into two datasets: training (DS1) and testing (DS2). This division is called the inter-patient paradigm, where the training and testing are performed in different sets of individuals. Table 2 shows the distribution of recordings across the datasets DS1 and DS2.

4. Methods

In the following, we present our proposed method for arrhythmia classification using single-lead ECG signals. Figure 1 shows all the necessary steps included in our methodology: filtering, segmentation, feature extraction, and classification.

4.1. Filtering

During ECG acquisition, the ECG signal is usually disturbed by different sorts of interference [18]. These interferences can damage the performance of classification systems. Therefore, a filtering step must be employed, as shown in Figure 1. In this work, the baseline wander was estimated using two moving average filters with 200 ms and 600 ms width, respectively. Then, the disturbance was subtracted from the input signal. Figure 2.a shows a segment of an ECG signal with baseline wander. The application of the two moving average filters results on the signal of the Figure 2.b. Finally, after

Table 1: Mapping between MIT-BIH and AAMI labels.

MIT-BIH class	AAMI class	Number of events
Normal beat (N or .) Left bundle branch block beat (L) Right bundle branch block beat (R) Atrial escape beat (e) Nodal (junctional) escape beat (j)	Normal (N)	90125
Atrial premature beat (A) Aberrated atrial premature beat (a) Nodal (junctional) premature beat (J) Supraventricular premature beat (S)	Supraventricular ectopic beat (S)	2781
Premature ventricular contraction (V) Ventricular escape beat (E)	Ventricular ectopic beat (V)	7009
Fusion of ventricular and normal beat (F)	Fusion beat (F)	803
Paced beat (P or /) Fusion of paced and normal beat (f) Unclassified beat (U)	Unknown beat (Q)	15

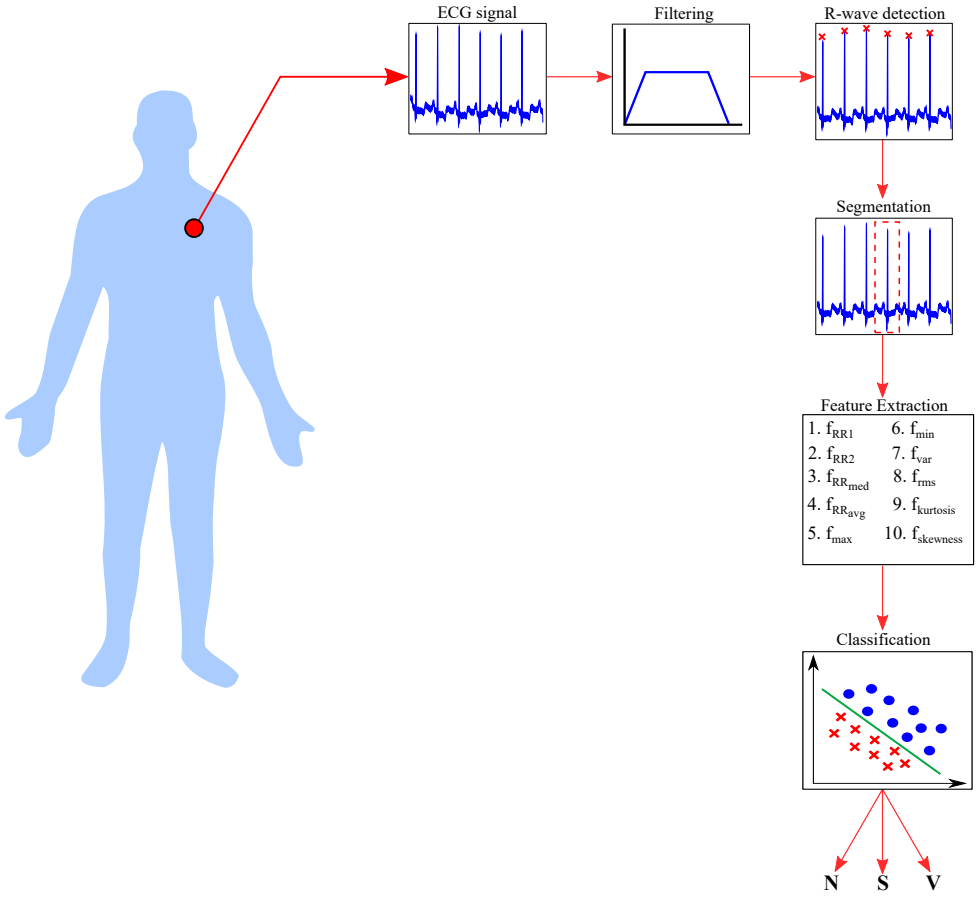

Figure 1: Proposed arrhythmia classification system.

Table 2: Distribution of the MIT-BIH recordings between training and testing.

Dataset	Recordings
DS1 (Training)	101, 106, 108, 109, 112, 114, 115, 116, 118, 119, 122, 124, 201, 203, 205, 207, 208, 209, 215, 220, 223, and 230.
DS2 (Testing)	100, 103, 105, 111, 113, 117, 121, 123, 200, 202, 210, 212, 213, 214, 219, 221, 222, 228, 231, 232, 233, and 234.

the subtraction, the denoised version of the ECG signal, shown in Figure 2.c, is obtained.

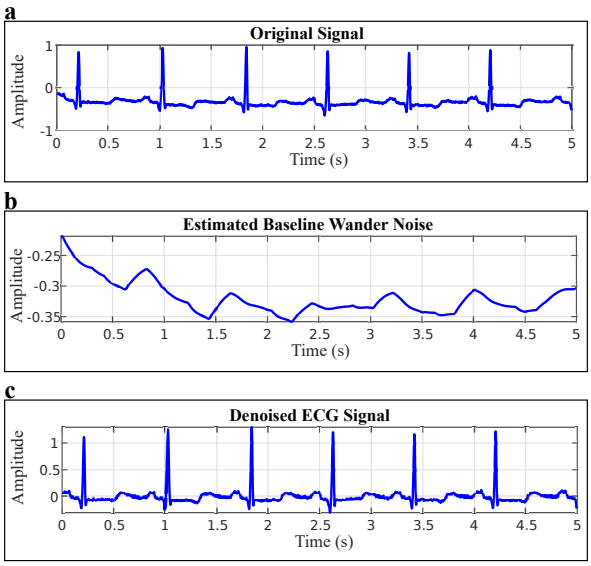

Figure 2: Proposed filtering process using two moving average filters.

4.2. Segmentation

After filtering, the R-waves are detected, and each heartbeat is segmented. Usually, the R-wave detection step is skipped in heartbeat classification papers, and the R-wave position given by the labeled database is used. However, in real scenarios, the detection of the R-wave is performed using algorithms, such as Pan-Tompkins [16]. Since these algorithms detect the R-wave position within an error, using the labeled R-wave position does not simulate a real scenario condition.

In this paper, we followed the suggestion of [7] and added jitter to the R-wave position given by the database. This jitter was modeled as a uniform distributed value between $-\delta$ and $+\delta$. The results of the proposed methodology are shown for $\delta = 0, 1, \ldots, 18$ samples. The maximum value of δ was chosen as 18 samples, i.e., 50 ms for a sampling rate of 360 Hz, because papers that analyze the performance of QRS detection algorithms consider errors up to 50 ms or 100 ms [14, 15]. Therefore, choosing $\delta = 18$ samples (100 ms tolerance window) provides a good simulation of a real scenario.

Afterward, given the R-wave position disturbed by the added jitter, we took 128 samples preceding and 127 succeeding the new R-wave position, i.e., 256 samples in total (including the R-wave sample). Figure 3 shows the segmentation process.

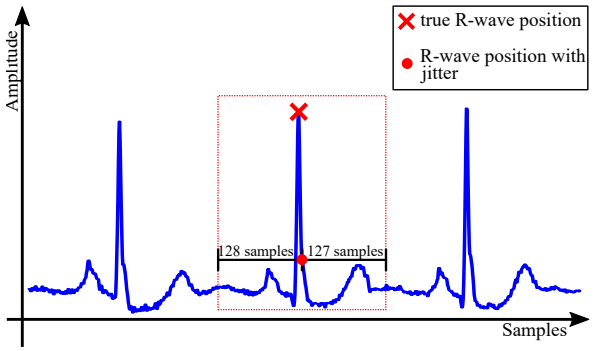

Figure 3: Segmentation process when the R-wave position is corrupted with jitter.

4.3. Feature Extraction

After the segmentation step, the necessary features to differentiate heartbeat classes are extracted from each heartbeat, as shown in Figure 1. Three groups of features are used for the proposed classification system: RR intervals, morphological, and high order statistics.

4.3.1. RR intervals

In this work, four features based on RR intervals are employed. They are described in the following.

Given the R-wave position of the i-th segmented ECG signal of the recording p, R(i), four RR features were used in the proposed classification system: f_{RR_1}, f_{RR_2}, $f_{RR_{med}}$, and $f_{RR_{avg}}$. They are defined in equations (1), (2), (3), and (4), respectively.

$$f_{RR_1} = \ln\left(RR_{-}(i)\right) \tag{1}$$

$$f_{RR_2} = \ln\left(RR_{+}(i)\right) \tag{2}$$

$$f_{RR_{med}} = \ln\Big(\text{median}\big(RR_{+}(i-15), \ldots, RR_{+}(i+15)\big)\Big) \tag{3}$$

$$\mathrm{f}_{\mathrm{RR}_{\mathrm{avg}}} = \ln\left(\frac{1}{\#\text{ segments in } p}\sum_{j=2}^{\#\text{ segments in } p}\mathrm{RR}_{-}(j)\right) \quad (4)$$

Where $\mathrm{RR(i)}_{+}$ and $\mathrm{RR(i)}_{-}$ are defined in the equations (5) and (6), respectively.

$$\mathrm{RR}_{+}(i) = \mathrm{R}(i+1) - \mathrm{R}(i) \quad (5)$$

$$\mathrm{RR}_{-}(i) = \mathrm{R}(i) - \mathrm{R}(i-1) \quad (6)$$

4.3.2. Morphological

Arrhythmia events can change the ECG morphology. To capture these changes, four morphological features were used in this study. They are described in the following.

Given a ECG segment $\boldsymbol{s}$ with $L = 256$ samples, the maximum ($\mathrm{f}_{\mathrm{max}}$), minimum ($\mathrm{f}_{\mathrm{min}}$), the variance ($\mathrm{f}_{\mathrm{var}}$) and the root-mean square ($\mathrm{f}_{\mathrm{rms}}$) values of $\boldsymbol{s}$ were employed as the morphological features. They are defined in equations (7), (8), (9), and (10), respectively.

$$\mathrm{f}_{\mathrm{max}} = \max\{s(i) : i = 1, \ldots, L\} \quad (7)$$

$$\mathrm{f}_{\mathrm{min}} = \min\{s(i) : i = 1, \ldots, L\} \quad (8)$$

$$\mathrm{f}_{\mathrm{var}} = \frac{1}{L-1}\sum_{i=1}^{L}(s(i) - \bar{s})^2 \quad (9)$$

$$\mathrm{f}_{\mathrm{rms}} = \sqrt{\frac{1}{L}\sum_{i=1}^{L}s(i)^2} \quad (10)$$

Where $\bar{s}$ is defined in the equation by

$$\bar{s} = \frac{1}{L}\sum_{i=1}^{L}s(i) \quad (11)$$

4.4. High order statistics

High order statistics (HOS) have been successively employed in many arrhythmia classification systems. HOS have a great capacity to capture signal nonlinearities. In this context, kurtosis ($\mathrm{f}_{\mathrm{kurtosis}}$) and skewness ($\mathrm{f}_{\mathrm{skewness}}$) were used as HOS features for the proposed classification system. They are defined in equations (12) and (13), respectively.

$$\mathrm{f}_{\mathrm{kurtosis}} = \frac{\frac{1}{L}\sum_{i=1}^{L}(s(i) - \bar{s})^4}{\left(\frac{1}{L}\sum_{i=1}^{L}(s(i) - \bar{s})^2\right)^2} \quad (12)$$

$$\mathrm{f}_{\mathrm{skewness}} = \frac{\frac{1}{L}\sum_{i=1}^{L}(s(i) - \bar{s})^3}{\left(\sqrt{\frac{1}{L}\sum_{i=1}^{L}(s(i) - \bar{s})^2}\right)^3} \quad (13)$$

4.5. Classification

Finally, a classifier must be employed to classify each heartbeat. This is the last step in our proposed arrhythmia classification system, as shown in Figure 1. The linear discriminant (LD) was the chosen algorithm for this work due to its simplicity and robustness against class imbalance. The LD classifier function for the class $\mathrm{k} \in \{\mathrm{N}, \mathrm{S}, \mathrm{V}\}$ is defined in the equation (14).

$$g_{\mathrm{k}}(\mathbf{x}) = \boldsymbol{\mu}_{\mathrm{k}}^{T}\boldsymbol{\Sigma}^{-1}\mathbf{x} - \frac{1}{2}\boldsymbol{\mu}_{\mathrm{k}}^{T}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_{\mathrm{k}} + \log(P(w_{\mathrm{k}})) \quad (14)$$

Where $\mathbf{x}$ is the feature vector, $\boldsymbol{\mu}_{\mathrm{k}}$ is the mean vector, $\boldsymbol{\Sigma}$ is the covariance matrix, and $P(\omega_{\mathrm{k}})$ is the prior probability. The values of $\boldsymbol{\mu}_{\mathrm{k}}$ and $\boldsymbol{\Sigma}$ are extracted from the training data using the equations (15) and (16), respectively.

$$\boldsymbol{\mu}_{\mathrm{k}} = \frac{1}{N_{\mathrm{k}}}\sum_{i=1}^{N_{\mathrm{k}}}\mathbf{x}_i \quad (15)$$

$$\boldsymbol{\Sigma} = \frac{1}{\sum_{\mathrm{k}=1}^{N_{\mathrm{classes}}} w_{\mathrm{k}}}\sum_{\mathrm{k}=1}^{N_{\mathrm{classes}}}\frac{w_{\mathrm{k}}\sum_{i=1}^{N_{\mathrm{k}}}(\mathbf{x}_i - \boldsymbol{\mu}_{\mathrm{k}})\cdot(\mathbf{x}_i - \boldsymbol{\mu}_{\mathrm{k}})^T}{N_{\mathrm{k}}} \quad (16)$$

Where N_{classes} is the number of classes, and N_{k} is the number of examples $\mathbf{x}_i$ of the class k. Besides, to compensate for the significant class imbalance of the MIT-BIH database, the prior probabilities ($P(w_{\mathrm{k}})$) were considered the same for all classes.

5. Results

5.1. Experimental Settings

The experiments were conducted in a 64-bit Windows 10 machine with 8 GB memory and i5-5200U 2.20 GHz in a MATLAB® environment. Our method was validated on the MIT-BIH database following the AAMI recommendations and the inter-patient paradigm. Besides, to demonstrate the robustness of our proposed classification system against segmentation errors, uniformly distributed jitter was added to the labeled R-wave position of the MIT-BIH database.

In order to evaluate the performance of the proposed classification system, three metrics were used: sensitivity (Se^{k}), positive predictive ($+\mathrm{P}^{\mathrm{k}}$), and F-score ($\mathrm{F}_{\mathrm{s}}^{\mathrm{k}}$).

Where k ∈ {N, S, V} indicates the corresponding class of each metric.

All of these metrics are based on measures of true positive (TP^k), false negative (FN^k), true negative (TN^k), and false positive (FP^k). The equations (17), (18), and (19) show the definition of the metrics Se^k, $+P^k$, and F_s^k, respectively.

$$Se^k = \frac{TP^k}{TP^k + FN^k} \times 100, \tag{17}$$

$$+P^k = \frac{TP^k}{TP^k + FP^k} \times 100, \tag{18}$$

$$F_s^k = 2\frac{Se^k(+P^k)}{Se^k + (+P^k)}. \tag{19}$$

5.2. Experimental Results

Since the jitter added to the labeled R-wave position is randomly generated, the experiments were repeated 33 times for each value of δ, and the final results are the average overall repetitions. The classification results for different values of δ are presented in Table 3. Also, Tables 4 and 5 show the obtained confusion matrix for $\delta = 0$ and $\delta = 18$, respectively. The former confusion matrix refers to one out of the 33 experiments that were generated.

For $\delta = 0$, our method achieves sensitivity, positive predictive value, and F-score of 94.5, 99.4, and 96.9 for the class N. For the class S, 92.5, 39.9, and 55.8 were achieved. Finally, class V obtained a sensitivity of 88.6, a positive predictive value of 94.6, and an F-score of 91.5.

When jitter is considered, i.e., $\delta = 18$, our method achieves 93.7, 89.7, 87.9 for the sensitivity of the classes, N, S, and V, respectively. Besides, the positive predictive values were 99.2, 36.8, and 93.9. Finally, the obtained F-score were 96.4, 52.2, and 90.8. It is possible to observe that even with $\delta = 18$, the system performance is not greatly impacted. Therefore, the proposed classification methodology is robust against jitter added to the R-wave position.

Additionally, we tested the robustness of each group of features against jitter. Tables 6, 7, and 8 presents the classification performance using only RR intervals, morphological values, and HOS features, respectively.

Table 9 compared our proposed methodology with $\delta = 0$ and $\delta = 18$ with other works in the literature that employed the inter-patient paradigm using three classes: N, S, and V. It is possible to observe in this table that our proposed method was able to outperform different state-of-the-art approaches even when jitter was considered.

In addition, we employed three alternative evaluation approaches for our proposed method. In these approaches, our method was evaluated with and without jitter. For the experiments related to the first approach, we consider DS2 dataset for training and the DS1 for testing, i.e., swap the setup proposed by [8]. Table 10 shows the results using this evaluation approach for $\delta = 0$ and $\delta = 18$. Also, Tables 11 and 12 show, respectively, the confusion matrix when $\delta = 0$ and $\delta = 18$ for this evaluation approach. Another investigated approach consists of employ a leave-one-patient-out cross-validation (LOPOCV). In this sort of evaluation, 43 out of 44 patients of the MIT-BIH dataset are used for training, and the testing is performed on the remaining patient.Hence, there are 44 training/testing procedures. Table 13 shows the obtained metrics for the whole MIT-BIH dataset using the LOPOCV approach. In addition, Tables 14 and 15 show, respectively, the confusion matrix for the LOPOCV approach for $\delta = 0$ and $\delta = 18$. Finally, as some works in the literature also consider the Fusion beat class (F), we also included the F class in our evaluation. Table 16 compared our proposed approach using four classes (N, S, V, and F) with other methods in the literature. In addition, Tables 17 and 18 show the confusion matrix for our method using four classes for $\delta = 0$ and $\delta = 18$, respectively.

Besides, we evaluated the time that our method needs to predict the class of a single heartbeat. On average, our method takes 2.6ms to evaluate each heartbeat. Considering that the heart beats every 333-1500ms (Heart rate: 40bpm-180bpm), the proposed approach can be employed in real-time.

6. Discussion

6.1. Robustness against segmentation error

As suggested by [7], we added jitter to the R-wave position given by the MIT-BIH database to demonstrate the robustness of the proposed system against segmentation errors. The same approach is carried out in [13]. This work uses a Gaussian distributed jitter with zero mean and 5 samples of standard deviation. However, this value does not comply with the expected R-wave detection errors generated by R-wave detection algorithms. Therefore, in our proposed method, we employed a maximum value of jitter of 18 samples, i.e., 50 ms. Therefore, by using the inter-patient paradigm and adding jitter to the R-wave position, our proposed experimental settings are closer to a real scenario than other works that do not follow the same procedures.

To evaluate the change in performance of the proposed classification when jitter increases, we applied an

Table 3: Classification performance with different values for the added jitter (δ).

δ	Class (N)			Class (S)			Class (V)		
	Se^N	$+P^N$	F_s^N	Se^S	$+P^S$	F_s^S	Se^V	$+P^V$	F_s^V
0	94.5	99.4	96.9	92.5	39.9	55.8	88.6	94.6	91.5
1	94.4	99.4	96.9	92.6	39.9	55.7	88.6	94.6	91.5
2	94.4	99.4	96.9	92.6	39.8	55.7	88.6	94.6	91.5
3	94.4	99.4	96.9	92.5	39.8	55.6	88.6	94.5	91.5
4	94.4	99.4	96.9	92.5	39.7	55.6	88.6	94.5	91.5
5	94.4	99.4	96.9	92.4	39.7	55.5	88.6	94.5	91.5
6	94.3	99.4	96.8	92.2	39.4	55.2	88.6	94.5	91.5
7	94.4	99.4	96.8	92.3	39.5	55.3	88.6	94.5	91.5
8	94.3	99.4	96.8	91.9	39.1	54.9	88.5	94.5	91.4
9	94.3	99.4	96.8	92.0	39.2	55.0	88.6	94.5	91.4
10	94.3	99.4	96.8	92.0	39.2	54.9	88.6	94.5	91.4
11	94.3	99.4	96.8	91.7	38.9	54.6	88.6	94.5	91.4
12	94.2	99.4	96.7	91.5	38.6	54.3	88.5	94.5	91.4
13	94.2	99.4	96.7	91.7	38.7	54.5	88.6	94.5	91.4
14	94.0	99.3	96.6	91.0	38.2	53.8	88.6	93.5	90.9
15	94.1	99.3	96.6	91.1	38.1	53.7	88.5	94.5	91.4
16	94.1	99.3	96.6	91.0	38.0	53.7	88.5	94.5	91.4
17	94.0	99.3	96.6	90.8	37.9	53.4	88.5	94.4	91.4
18	93.7	99.2	96.4	89.7	36.8	52.2	87.9	93.9	90.8

Table 4: Confusion matrix for the proposed classification system for $\delta = 0$

		Predicted			
		N	S	V	Total
Truth	N	41071	2265	145	43481
	S	125	1673	10	1808
	V	108	254	2818	3180
	Total	41304	4192	2973	

Table 5: Confusion matrix for the proposed classification system for $\delta = 18$

		Predicted			
		N	S	V	Total
Truth	N	40873	2463	145	43481
	S	178	1619	11	1808
	V	121	265	2794	3180
	Total	41172	4347	2950	

increasing amount of jitter ($\delta = 0, 1, 2, \ldots, 18$). Since the jitter is uniformly distributed, the experiments were conducted 33 times, and the average overall experiments are displayed in Table 3. Looking at this table, it is possible to observe that the performance of our proposed system does not vary in a significant way. For classes N and V, for example, the greatest found variation were 0.8% for both Se^N and Se^V, respectively. On the other hand, class S is the one that is more jeopardized by segmentation errors. The value of $+P^S$, for example, varies 7.8% when δ changes from 0 to 18. Therefore, the proposed system is robust against segmentation errors.

6.2. Robustness of each group of features against segmentation error

In this section, we also analyze the performance difference caused by jitter on each individual group of features.

In Table 6, we considered only the RR intervals for classification purposes. It is possible to notice a small performance decrease for the class N, where Se^N and $+P^N$ decreased their performance by 2.6% and 0.2%, respectively. For the class S, as the jitter increases, there is a trade-off between sensitivity and the positive predictive value. The former, Se^S, increases by 1.3%, but the latter, $+P^S$, decreases by 6.2%. Besides, the sensitivity and positive predictive value of the class V (Se^V and $+P^V$) decreased by 2.0% and 16.8%, respectively. Thus it is possible to conclude that RR features present good robustness against jitter for the classes N and S. On the other hand, for the class V, the performance is significantly deteriorated.

Table 6: Classification performance for RR features with different values for the added jitter (δ).

δ	Class (N)			Class (S)			Class (V)		
	Se^N	$+P^N$	F_s^N	Se^S	$+P^S$	F_s^S	Se^V	$+P^V$	F_s^V
0	93.3	99.0	96.1	68.6	37.1	48.1	81.7	63.0	71.1
1	93.3	99.0	96.1	68.7	37.0	48.1	81.6	63.0	71.1
2	93.3	99.0	96.1	68.7	36.9	48.0	81.6	63.0	71.1
3	93.3	99.0	96.1	68.7	36.9	48.0	81.6	63.0	71.1
4	93.3	99.0	96.1	68.7	36.9	48.0	81.5	62.8	71.0
5	93.3	99.0	96.0	69.2	37.0	48.2	81.3	62.9	70.9
6	93.2	99.0	96.0	68.9	36.8	48.0	81.4	62.7	70.8
7	93.2	99.0	96.0	68.7	36.7	47.8	81.3	62.4	70.6
8	93.2	99.0	96.0	68.9	36.7	47.9	81.2	62.3	70.5
9	93.1	99.0	95.9	68.9	36.5	47.7	81.1	61.9	70.2
10	92.9	98.9	95.8	69.3	36.3	47.7	80.9	61.3	69.7
11	92.8	98.9	95.8	69.2	36.3	47.6	80.8	60.7	69.3
12	92.7	98.9	95.7	69.0	36.1	47.4	80.7	60.0	68.8
13	92.5	98.9	95.6	69.1	36.0	47.3	80.5	59.3	68.3
14	92.3	98.9	95.5	69.5	36.0	47.4	80.3	58.2	67.5
15	92.0	98.9	95.3	68.6	35.4	46.7	80.5	56.7	66.5
16	91.7	98.8	95.1	69.9	35.5	47.1	80.0	55.4	65.4
17	91.6	98.9	95.1	69.2	35.4	46.8	80.0	54.5	64.8
18	90.9	98.8	94.7	69.5	34.8	46.4	80.1	52.4	63.3

Table 7: Classification performance for morphological features with different values for the added jitter (δ).

δ	Class (N)			Class (S)			Class (V)		
	Se^N	$+P^N$	F_s^N	Se^S	$+P^S$	F_s^S	Se^V	$+P^V$	F_s^V
0	79.9	98.3	88.2	79.1	14.0	23.7	83.1	92.4	87.5
1	80.0	98.3	88.2	79.1	14.0	23.8	83.2	92.4	87.5
2	80.1	98.3	88.3	79.1	14.1	23.9	83.2	92.4	87.5
3	80.3	98.3	88.4	79.0	14.2	24.0	83.2	92.4	87.5
4	80.6	98.3	88.6	78.8	14.2	24.1	83.1	92.4	87.5
5	80.9	98.3	88.7	78.4	14.4	24.3	83.1	92.4	87.5
6	81.1	98.3	88.9	78.1	14.5	24.4	83.1	92.4	87.5
7	81.5	98.3	89.1	77.7	14.7	24.7	83.0	92.4	87.5
8	81.7	98.3	89.3	77.5	14.8	24.8	82.9	92.5	87.4
9	82.2	98.3	89.5	77.0	15.0	25.1	82.9	92.5	87.4
10	82.5	98.3	89.7	76.6	15.1	25.3	82.8	92.4	87.4
11	82.8	98.3	89.9	76.1	15.3	25.5	82.7	92.4	87.3
12	83.2	98.3	90.1	75.5	15.5	25.7	82.6	92.4	87.2
13	83.5	98.3	90.3	74.9	15.6	25.8	82.5	92.3	87.2
14	84.0	98.2	90.6	73.4	15.7	25.8	82.4	92.3	87.1
15	84.1	98.2	90.6	72.6	15.7	25.7	82.4	92.2	87.0
16	84.4	98.2	90.8	71.1	15.6	25.6	82.4	92.1	87.0
17	84.6	98.1	90.9	69.6	15.5	25.3	82.2	92.0	86.8
18	84.9	98.1	91.0	68.0	15.4	25.1	82.1	91.9	86.7

Table 7 shows the classification performance when only morphological features are employed. For the class N, there was a 6.3% increase on the sensitivity (Se^N), but the positive predictive value, $+P^N$, stayed almost steady (decreased by only 0.2%). This result shows that there was a decrease in the false negatives for the

Table 8: Classification performance for HOS features with different values for the added jitter (δ).

δ	Class (N)			Class (S)			Class (V)		
	Se^N	$+P^N$	F_s^N	Se^S	$+P^S$	F_s^S	Se^V	$+P^V$	F_s^V
0	73.0	98.5	83.9	2.2	1.5	1.7	71.1	16.7	27.0
1	73.0	98.5	83.9	2.4	1.6	1.9	71.1	16.7	27.0
2	73.0	98.5	83.9	2.6	1.8	2.1	71.1	16.7	27.0
3	73.0	98.5	83.8	2.9	2.0	2.3	71.1	16.7	27.0
4	72.9	98.5	83.8	3.4	2.3	2.7	71.1	16.7	27.0
5	72.8	98.6	83.8	4.0	2.6	3.2	71.0	16.6	27.0
6	72.7	98.6	83.7	4.9	3.2	3.8	71.1	16.6	26.9
7	72.6	98.6	83.7	6.0	3.8	4.6	71.1	16.6	26.9
8	72.5	98.6	83.6	7.1	4.4	5.4	71.0	16.6	27.0
9	72.4	98.7	83.5	8.3	4.9	6.2	71.1	16.7	27.0
10	72.2	98.7	83.4	9.1	5.3	6.7	71.1	16.7	27.0
11	72.0	98.7	83.2	10.4	5.8	7.4	71.2	16.7	27.1
12	71.8	98.7	83.1	11.4	6.1	7.9	71.1	16.8	27.1
13	71.5	98.7	83.0	12.2	6.2	8.3	71.1	16.8	27.2
14	71.3	98.8	82.8	13.3	6.5	8.8	71.2	16.9	27.3
15	71.1	98.8	82.7	14.1	6.7	9.1	71.2	17.0	27.4
16	70.8	98.8	82.5	15.2	6.9	9.4	71.3	17.0	27.5
17	70.4	98.8	82.2	16.3	6.9	9.7	71.4	17.1	27.6
18	69.8	98.8	81.8	17.0	6.7	9.6	71.4	17.3	27.8

Table 9: Comparison between the proposed methodology and different state-of-the-art methods.

Work	Class (N)			Class (S)			Class (V)		
	Se^N	$+P^N$	F_s^N	Se^S	$+P^S$	F_s^S	Se^V	$+P^V$	F_s^V
Proposed (δ = 0)	94.5	**99.4**	**96.9**	**92.5**	39.9	55.8	**88.6**	**94.6**	**91.5**
Proposed (δ = 18)	93.7	99.2	96.4	89.7	36.8	52.2	87.9	93.9	90.8
Lin and Yang (2014) [12]	91.6	99.3	95.3	81.4	31.6	45.5	86.2	73.7	79.5
Garcia et. al (2016) [31]	**95.0**	96.5	95.7	29.6	26.4	27.9	85.1	66.3	74.5
Garcia et. al (2017) [32]	94.0	98.0	96.0	62.0	**53.0**	**57.1**	87.3	59.4	70.7

* Values in bold indicate the best result.

Table 10: Classification perfomance using DS2 for training and DS1 for testing.

δ	Class (N)			Class (S)			Class (V)		
	Se^N	$+P^N$	F_s^N	Se^S	$+P^S$	F_s^S	Se^V	$+P^V$	F_s^V
0	90.9	97.2	93.9	41.4	13.2	20	72.7	58	64.5
18	90.5	97.1	93.7	42.1	12.8	19.7	72.5	57.7	64.3

class N. On the other hand, for the classes S, the sensitivity (Se^S) decreased by 14.0%. The positive predictive values ($+P^S$) increased by 10.0%. Both the sensitivity (Se^V) and positive predictive value ($+P^V$) of the class V decreased. The former decreased by 1.2%, and the latter decreased by 0.5%. Therefore, as jitter increases, it is possible to realize that the class imbalance, an already known problem for the MIT-BIH dataset, becomes worse. The classifier tends to choose the class N despite the others.

The results for the classification system when only HOS features are considered are displayed in Table 8. In this table, we can see a decrease in the class N's sensitivity (4.4%). The positive predictive value, on the other hand, stayed almost still (increase by 0.3%). For class S, there is a great increase in both sensitivity and positive predictive value. They raised by 672.7% and 346.7%, respectively. The sensitivity of the class V increased by only 0.4%, and its positive predictive value increased by 3.6%. From these results, we can say that with in-

Table 11: Confusion matrix using DS2 for training and DS1 for testing ($\delta = 0$).

		Predicted			
		N	S	V	Total
Truth	N	40983	2236	1882	45101
	S	446	379	90	915
	V	756	264	2720	3740
	Total	42185	2879	4692	

Table 12: Confusion matrix using DS2 for training and DS1 for testing ($\delta = 18$).

		Predicted			
		N	S	V	Total
Truth	N	40846	2360	1895	45101
	S	450	376	89	915
	V	791	242	2707	3740
	Total	42087	2978	4691	

creasing jitter, the HOS based classifier does not have a considerable effect on the metrics of the class V. On the other hand, the decrease on Se^N and rise of both Se^S and $+P^S$ indicates that the classifier, with jitter, is increasing its bias towards class S over class N. It is important to say that besides the huge percentage increases in class S's metrics, their values are still considerably low.

6.3. Comparison with other works

Considering our proposed system with no jitter, i.e., $\delta = 0$, it is possible to observe in Table 9 that our method overcomes [12], [31], and [32] for the metrics $+P^N$, F_s^N, Se^S, Se^V, $+P^V$, and F_s^V. Additionally, for Se^N, our result is superior to [12] and [32], but it is inferior to [31] by 0.5%. Besides having the highest sensitivity for the class S, our method has a low $+P^S$. It is inferior to [32] by 32.8%. Consequently, the F_s^S value is also inferior to [32].

With the maximum amount of jitter, i.e., $\delta = 18$, our proposed approach also presents a good performance compared to other state-of-the-art methods. It has superior results in comparison to [12], [31], and [32] for the metrics F_s^N, Se^S, Se^V, $+P^V$, and F_s^V. In addition, for the metrics Se^N and $+P^N$, the performance of our method is inferior by only 1.4% and 0.1%, respectively, in comparison to the best results for these metrics.

6.4. Alternative evaluation approaches

6.4.1. Employing DS2 as the training set

As can be seen in Table 10, in this alternative evaluation approach, there is a considerable decrease regarding the metrics in comparison with the Chazal et al. [8] evaluation approach, especially for class S. For the metrics Se^N, $+P^N$, F_s^N, Se^S, $+P^S$, F_s^S, Se^V, $+P^V$, and F_s^V , our method achieves 90.9, 97.2, 93.9, 41.4, 13.2, 20.0, 72.7, 58.0, and 64.5, respectively, when no jitter is considered. When a jitter of $\delta = 18$ is employed, our method achieves 90.5, 97.1, 93.7, 42.1, 12.8, 19.7, 72.5, 57.7, and 64.3.

The obtained results can be explained by the high concentration of S class heartbeats on a few patients of the DS2 dataset. Patient 232 (from MIT-BIH DB) , for example, concentrates 75% of heartbeats from the S class. This unbalanced heartbeat distribution can lead to poor generalization. Therefore, to avoid choosing non-representative training datasets, a leave-one-patient-out cross-validation approach is better suited for evaluation methods for the MIT-BIH dataset since this dataset only contains a few patients.

6.4.2. Leave-one-out cross-validation

In Table 13, we displayed the obtained metrics for our method using the LOPOCV approach for $\delta = 0$ and $\delta = 18$. It is possible to observe a decrease in some metrics, if we compare the LOPOCV evaluation with the evaluation proposed by de Chazal et al. [8]. However, this decrease is not as significant as reported in the former alternative evaluation approach (DS2 for training and DS1 for testing). We also emphasize that our method still shows robustness against added jitter in this evaluation approach. The highest obtained variation was 7.4% for the positive predictive value of the class S ($+P^V$).

In addition, since the MIT-BIH dataset presents only a few patients available, we consider that the LOPOCV approach is the fairest evaluation approach for this dataset. Using DS1 as training and DS2 as testing or the opposite can lead to skewed data distribution that does not generalize well to real scenarios. Hence, we suggest that future works in the literature also employ this evaluation approach to better compare methods.

6.4.3. Including the Fusion beat class

In Table 16, we showed our proposed method using $\delta = 0$ and $\delta = 18$, considering 4 classes: N, S, V, and F. In this table, we also displayed other works in the literature that used the Fusion beat class.

It can be seen that the addition of a new class to the problem (class F) caused a significant decrease in the sensitivity for class N. By analyzing the confusion matrix in Tables 17 and 18, we can see that a considerable amount of N examples are classified as F for both $\delta = 0$ and $\delta = 18$. This leads to a decrease in the N class's sen-

Table 13: Classification performance using LOPOCV approach.

δ	Class (N)			Class (S)			Class (V)		
	$\mathbf{Se^N}$	$\mathbf{+P^N}$	$\mathbf{F_s^N}$	$\mathbf{Se^S}$	$\mathbf{+P^S}$	$\mathbf{F_s^S}$	$\mathbf{Se^V}$	$\mathbf{+P^V}$	$\mathbf{F_s^V}$
0	91.7	98.2	94.8	73.5	25.5	37.8	78.5	71.6	75.4
18	91.0	98.0	94.4	72.1	23.6	35.6	78.7	70.8	74.5

Table 14: Confusion matrix using LOPOCV approach ($\delta = 0$).

		Predicted			
		N	S	V	Total
Truth	N	81203	5295	2084	88582
	S	627	2001	95	2723
	V	859	559	5502	6920
	Total	82689	7855	7681	

Table 15: Confusion matrix using LOPOCV approach ($\delta = 18$).

		Predicted			
		N	S	V	Total
Truth	N	80611	5826	2145	88582
	S	662	1964	97	2723
	V	945	532	5443	6920
	Total	82218	8322	7685	

sitivity and poor performance for the F class's positive predictive value.

In comparison with other works in the literature that also considered the F class [8, 11, 9], our method with $\delta = 0$ achieves superior results for the metrics $+P^N$, Se^S and $+P^V$. When we employ $\delta = 18$, our method also achieves competitive results compared to the literature. Besides not having the best results for the metrics $+P^N$, Se^S, $+P^V$, and F_s^V, our method is inferior to the best result by only 0.1%, 1.2%, 0.1%, and 1.9%, respectively.

7. Conclusion

In this work, we proposed an arrhythmia classification system using single-lead ECG signals. Initially, a filtering process was applied, combining two moving average filters to decrease baseline wander noise. Then, the heartbeats were segmented using the position of the R-wave given by the MIT-BIH database added with an artificially generated jitter. Also, a combination of features based on RR intervals, morphological values, and higher-order statistics was extracted from each beat. Finally, the extracted features were used to classify arrhythmias using a linear discriminant classifier. To demonstrate the applicability of the proposed method in real situations, the inter-patient paradigm was employed using the MIT-BIH database. Additionally, the classification performance of each group of features was analyzed separately in order to access the robustness of each one against jitter.

The results demonstrated that the proposed method achieves competitive results when comparing with other state-of-the-art techniques, even with added jitter to the R-wave position. Thus, it was possible to demonstrate that our approach is feasible to be applied in real scenarios where noise, errors in the detection of the R-wave, and the inter-patient paradigm are present.

Acknowledgements

The authors would like to thank the support of the Federal University of Juiz de Fora, FAPEMIG, and INESC.

References

[1] M. AlGhatrif, J. Lindsay, A brief review: history to understand fundamentals of electrocardiography, Journal of community hospital internal medicine perspectives 2 (1) (2012) 14383.

[2] World Health Organization, Cardiovascular disease. URL `https://www.who.int/cardiovascular_diseases/en/`

[3] T. A. Gaziano, A. Bitton, S. Anand, S. Abrahams-Gessel, A. Murphy, Growing epidemic of coronary heart disease in low- and middle-income countries, Current problems in cardiology 35 (2) (2010) 72–115.

[4] American Heart Association, Holter monitor. URL `https://www.heart.org/en/health-topics/heart-attack/diagnosing-a-heart-attack/holter-monitor`

[5] İ. Güler, E. D. Übeylı, Ecg beat classifier designed by combined neural network model, Pattern recognition 38 (2) (2005) 199–208.

[6] M. H. Song, J. Lee, S. P. Cho, K. J. Lee, S. K. Yoo, Support vector machine based arrhythmia classification using reduced features, International Journal of Control, Automation, and Systems 3 (4) (2005) 571–579.

[7] E. J. d. S. Luz, W. R. Schwartz, G. Cámara-Chávez, D. Menotti, Ecg-based heartbeat classification for arrhythmia detection: A survey, Computer methods and programs in biomedicine 127 (2016) 144–164.

[8] P. De Chazal, M. O'Dwyer, R. B. Reilly, Automatic classification of heartbeats using ecg morphology and heartbeat interval features, IEEE transactions on biomedical engineering 51 (7) (2004) 1196–1206.

Table 16: Comparison between the proposed methodology and different state-of-the-art methods that uses four classes: N, S, V, and F.

Work	Class (N)			Class (S)			Class (V)			Class (F)		
	Se^N	$+P^N$	F_s^N	Se^S	$+P^S$	F_s^S	Se^V	$+P^V$	F_s^V	Se^F	$+P^F$	F_s^F
Proposed (δ = 0)	79.2	**99.6**	88.2	**92.2**	39.7	55.0	87.2	**92.8**	89.9	81.4	4.6	8.6
Proposed (δ = 18)	77.9	99.4	87.3	90.6	36.9	52.5	87.2	92.6	89.8	81.4	4.4	8.3
De Chazal et al. (2004) [8]	86.9	99.2	92.6	75.9	38.5	51.1	77.7	81.6	79.6	89.4	8.6	15.7
Zhang et al. (2014) [11]	88.9	99.0	93.7	79.1	36.0	49.5	85.5	92.7	89.0	**93.8**	13.7	23.9
Shi et al. (2019) [9]	**92.1**	99.5	**95.7**	91.7	**46.2**	**61.4**	**95.1**	88.1	**91.5**	61.6	**15.2**	**24.4**

* Values in bold indicate the best result.

Table 17: Confusion matrix for the proposed method using four classes: N, S, V, and F (δ = 0).

		Predicted				
		N	S	V	F	Total
Truth	N	34629	2217	140	6495	43481
	S	98	1651	8	51	1808
	V	58	224	2775	123	3180
	F	16	1	56	314	387
	Total	34801	4093	2979	6983	

Table 18: Confusion matrix for the proposed method using four classes: N, S, V, and F (δ = 18).

		Predicted				
		N	S	V	F	Total
Truth	N	34216	2432	142	6691	43481
	S	141	1616	9	42	1808
	V	81	255	2752	92	3180
	F	15	1	59	312	387
	Total	34453	4304	2962	7137	

[9] H. Shi, H. Wang, Y. Huang, L. Zhao, C. Qin, C. Liu, A hierarchical method based on weighted extreme gradient boosting in ecg heartbeat classification, Computer methods and programs in biomedicine 171 (2019) 1–10.

[10] S. Raj, K. C. Ray, Automated recognition of cardiac arrhythmias using sparse decomposition over composite dictionary, Computer methods and programs in biomedicine 165 (2018) 175–186.

[11] Z. Zhang, J. Dong, X. Luo, K.-S. Choi, X. Wu, Heartbeat classification using disease-specific feature selection, Computers in biology and medicine 46 (2014) 79–89.

[12] C.-C. Lin, C.-M. Yang, Heartbeat classification using normalized rr intervals and morphological features, Mathematical Problems in Engineering 2014.

[13] C. Ye, B. V. Kumar, M. T. Coimbra, Heartbeat classification using morphological and dynamic features of ecg signals, IEEE Transactions on Biomedical Engineering 59 (10) (2012) 2930–2941.

[14] F. Liu, C. Liu, X. Jiang, Z. Zhang, Y. Zhang, J. Li, S. Wei, Performance analysis of ten common qrs detectors on different ecg application cases, Journal of healthcare engineering 2018.

[15] Z. Zidelmal, A. Amirou, M. Adnane, A. Belouchrani, Qrs detection based on wavelet coefficients, Computer methods and programs in biomedicine 107 (3) (2012) 490–496.

[16] J. Pan, W. J. Tompkins, A real-time qrs detection algorithm, IEEE transactions on biomedical engineering (3) (1985) 230–236.

[17] G. B. Moody, R. G. Mark, The mit-bih arrhythmia database on cd-rom and software for use with it, in: [1990] Proceedings Computers in Cardiology, IEEE, 1990, pp. 185–188.

[18] G. D. Clifford, F. Azuaje, P. Mcsharry, Ecg statistics, noise, artifacts, and missing data, Advanced methods and tools for ECG data analysis 6 (2006) 18.

[19] R. Sameni, M. Shamsollahi, C. Jutten, M. Babaie-Zade, Filtering noisy ecg signals using the extended kalman filter based on a modified dynamic ecg model, in: Computers in Cardiology, 2005, IEEE, 2005, pp. 1017–1020.

[20] V. Queiroz, E. Luz, G. Moreira, Á. Guarda, D. Menotti, Automatic cardiac arrhythmia detection and classification using vectorcardiograms and complex networks, in: 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2015, pp. 5203–5206.

[21] M. Alfaouri, K. Daqrouq, Ecg signal denoising by wavelet transform thresholding, American Journal of applied sciences 5 (3) (2008) 276–281.

[22] N. Nikolaev, Z. Nikolov, A. Gotchev, K. Egiazarian, Wavelet domain wiener filtering for ecg denoising using improved signal estimate, in: 2000 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No. 00CH37100), Vol. 6, IEEE, 2000, pp. 3578–3581.

[23] B. N. Singh, A. K. Tiwari, Optimal selection of wavelet basis function applied to ecg signal denoising, Digital signal processing 16 (3) (2006) 275–287.

[24] M. Blanco-Velasco, B. Weng, K. E. Barner, Ecg signal denoising and baseline wander correction based on the empirical mode decomposition, Computers in biology and medicine 38 (1) (2008) 1–13.

[25] Z. Wang, F. Wan, C. M. Wong, L. Zhang, Adaptive fourier decomposition based ecg denoising, Computers in Biology and Medicine 77 (2016) 195–205.

[26] Y. Wu, L. Zhang, Ecg classification using ica features and support vector machines, in: International Conference on Neural Information Processing, Springer, 2011, pp. 146–154.

[27] M. Sarfraz, A. A. Khan, F. F. Li, Using independent component analysis to obtain feature space for reliable ecg arrhythmia classification, in: 2014 IEEE international conference on bioinformatics and biomedicine (BIBM), IEEE, 2014, pp. 62–67.

[28] Y. Kutlu, D. Kuntalp, Feature extraction for ecg heartbeats using higher order statistics of wpd coefficients, Computer methods and programs in biomedicine 105 (3) (2012) 257–267.

[29] G. Garcia, G. Moreira, D. Menotti, E. Luz, Inter-patient ecg heartbeat classification with temporal vcg optimized by pso, Scientific reports 7 (1) (2017) 10543.

[30] E. J. D. S. Luz, T. M. Nunes, V. H. C. De Albuquerque,

J. P. Papa, D. Menotti, Ecg arrhythmia classification based on optimum-path forest, Expert Systems with Applications 40 (9) (2013) 3561–3573.

[31] G. Garcia, G. Moreira, E. Luz, D. Menotti, Improving automatic cardiac arrhythmia classification: Joining temporal-vcg, complex networks and svm classifier, in: 2016 International Joint Conference on Neural Networks (IJCNN), IEEE, 2016, pp. 3896–3900.

[32] G. Garcia, G. Moreira, D. Menotti, E. Luz, Inter-patient ecg heartbeat classification with temporal vcg optimized by pso, Scientific Reports 7 (1) (2017) 1–11.