

Classifying cardiac rhythms by means of digital signal processing and machine learning

N. Pinho, D. Azevedo, and A. Santos

Abstract—Electrocardiogram (ECG) measures the electrical activity of the heart, which can be used in the diagnosis of different heart diseases. In the scientific literature there are many studies that have been applied machine learning for recognizing ECG patterns, where most of them attempt to classify heart beats. This paper presents a novel methodology for automatically classifying seventeen cardiac rhythms by means of digital signal processing and machine learning. The steps before the classification include the mapping of ECG signal to the frequency domain through power spectrum density, class balance with Adaptive Synthetic Sampling algorithm, and statistical normalization. The classifiers employed were Support Vector Machine, Multilayer Perceptron Neural Network, k-Nearest Neighbors, and Random Forest. The results showed accuracy, sensitivity, specificity, and Fleiss' kappa of up to 98.86%, 99.93%, 98.85%, and 89.68%, respectively, which are relatively better than the performance observed in the state-of-the-art works. In addition, this study highlighted that when the class balance procedure is applied, the classification step becomes less complex and can increase in terms of performance.

Index Terms—ECG, Digital signal processing, Machine learning, Cardiac rhythm

I. INTRODUCTION

ACCORDING to the World Health Organization, the cardiovascular diseases are the main cause of death in the world. These diseases caused the death of approximately 17.9 million people by year, i.e., about 31% of all deaths in the world [1]. Incipient diagnosis is essential in the treatment of these health problems, and the electrocardiogram (ECG) is an effective tool in this context, which is also crucial in the normal cardiac rhythm management. ECG is used in the diagnosis of the cause of chest pain and suitable early intervention in the myocardial infarction. It is a measure of the electric pulse generated in the heart which propagates to the surface of the skin, being its evaluation carried out based on pattern recognition [2]. In general, the classification of an ECG signal is accomplished by a cardiologist, but this task might be slow and expensive. As an alternative, in the last decade has growth the research of algorithms to automatize the ECG evaluation, mainly by the application of methods related to the Artificial Intelligence (AI) field [3]–[5].

In the AI field there is a research area called machine learning, which aims to make a machine (or algorithm) to detect and extrapolate new patterns, with the capability of adapting to new circumstances. This area can be divided in three branches: supervised, unsupervised, and reinforcement

learning. The problem tackled in supervised learning is focused on learning a function from examples in its inputs and outputs, while the unsupervised learning aims to learn patterns in its inputs, but no outputs are specified. In the reinforcement learning, an agent must learn to perform some task according to the reward related to an input [6]. The classification of ECG signals is mostly performed in the context of a supervised learning problem.

In the literature, there are many works related to pattern recognition on ECG signals, in which the heart beat type recognition was explored, as in [7]–[9], for instance. There are many approaches employed to identify the heart beats, but the general methodology is: (1) Acquiring the database; (2) Preprocessing the ECG signals; (3) Detecting the QRS structure (the region around the periodic peaks in ECG signal); (4) Performing signal segmentation; (5) Extracting features; (6) Training the classifier; (7) Testing the classifier; and (8) Evaluating and analyzing the results.

A methodology for classifying Normal Sinus Rhythm (N) and heart arrhythmia by means of statistics in time domain was presented in [3]. The methodology was divided in seven steps: (1) database reorganization; (2) noise elimination; (3) detection of the QRS structure; (4) computation of statistics; (5) anomaly detection, (6) anomaly type identification; and (7) estimation of the average accuracy. This approach yielded an accuracy of 98.78%, but only a classifier and a single metric to evaluate the classification performance were employed.

The classification of distinct heart rhythms was performed in [5], such as: N, Atrial Fibrillation (AFIB), Acute Myocardial Infarction (MI), and Congestive Heart Failure (CHF). For such a task, in the first step, the time series based on RR intervals (interval between R waves) of N, AFIB, MI, and CHF were obtained from the Physionet database [10]. In the second step, different characteristics from Poincare plot were extracted for Heart Rate Variability Analysis. In the last step, these features were used as input to a k-Nearest Neighbors (KNN) classifier. An approach with more classifiers would be more interesting for a comparison study. The results presented for the test stage were sensitivity, specificity, and accuracy equal to 94.64%, 98.21%, and 97.31%, respectively.

A robust methodology was proposed in [4] for recognizing cardiac health, which was based on the steps presented previously, and evaluated the classifiers using the raw signal, and scaled or normalized versions of the former. Support Vector Machine (SVM), KNN, Probabilistic Neural Network, and Radial Basis Function Neural Network were employed for the classification task, in which the performance was evaluated via nine metrics, and the number of classes evaluated was seventeen. This approach yielded accuracy close to 99%, but

N. Pinho, D. Azevedo, and A. Santos are with Faculty of Computing and Electrical Engineering, Universidade Federal do Sul e Sudeste do Pará, F. 17, Q. 4, L. E., Marabá, Pará 68505-080, Brazil (e-mail: nielson-silva@unifesspa.edu.br).

Digital Object Identifier: 10.14209/jcis.2020.3

the computational cost was high, mainly due to the use of a genetic algorithm to adjust the parameters of the classifiers, which required between 135 and 170 hours. Another drawback was the very unbalanced number of segments per class, which brought greater complexity to the searching for the optimal parameters of each classifier.

This work presents a novel methodology to classify distinct cardiac rhythms by applying the following steps: (1) segmentation, (2) class balance with Adaptive Synthetic Sampling (ADASYN), (3) statistical normalization, (4) feature extraction, (5) training, optimizing, and testing of classifiers, and (6) evaluation of results. The applied classifiers are SVM, KNN, Random Forest (RF), and Multi Layer Perceptron (MLP). Besides, to evaluate the classification performance, four metrics is taking into account, such as accuracy, sensitivity, specificity, and Fleiss' kappa. This methodology is less computationally expensive and more simple when compared with the work in [4], because the class balance procedure attenuates the computational complexity required for the following steps, without compromises the classification performance.

The remainder of this paper is organized as follows. In Section II, the ECG signal, its main characteristics and how it can be used to recognize patterns related to cardiac diseases are highlighted. Section III comprises the aspects of the database used in this study, as well as the parameters applied in the ECG data acquisition. The whole methodology of this work is described in more detail in Sections IV, V, and VI. The obtained results, through the application of the proposed methodology on the database that encompass seventeen different cardiac rhythms, are shown and discussed in Section VII. Finally, Section VIII synthesizes the main strengths and challenges of this methodology.

II. ELECTROCARDIOGRAM

In this section the ECG signal, its main characteristics and how it can be used to identify cardiac diseases are presented.

When the cardiac pulse passes through the heart, an electric current flows from the heart to adjacent tissues, and a part of this signal achieves the skin surface. The electric potential of this signal can be measured with electrodes attached in skin regions in which this pulse appears; the reason to call this signal as electrocardiogram [11]. This signal is low frequency, in general ranging from 0.05 Hz to 100 Hz, with the major part of the energy up to 35 Hz, and its amplitude varies from 10 μV to 5 mV, being the typical value around 1 mV [12]. ECG has a standard structure with peaks and valleys, labeled by the letters P, Q, R, S, and T. Fig. 1 depicts the typical appearance of this structure.

In the normal ECG of an adult person, the P wave has period from 0.12 s to 0.13 s; the PR interval, measured from the begin of P wave up to the begin of the QRS complex, has period from 0.12 s up to 0.20 s; and the QRS complex goes from 0.07 s up to 0.10 s [14].

When a person is in normal rhythm, the P wave is a positive deflection, the Q wave is a negative deflection followed by the R wave with a relatively high positive amplitude and S is a small negative deflection. After the ST interval, a positive

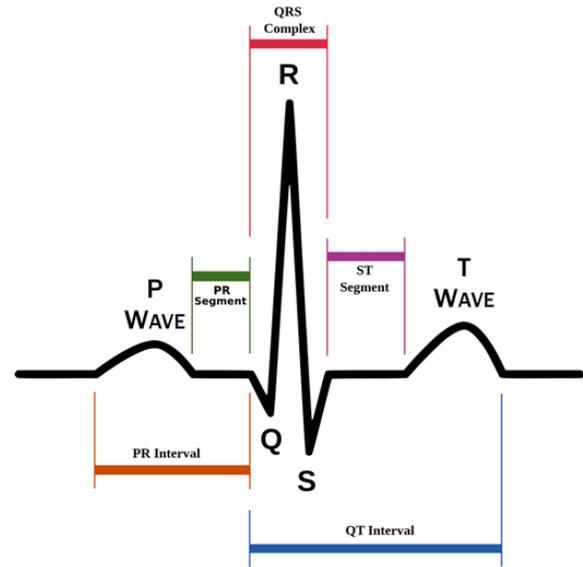


Figure 1. Standard structure of an ECG signal [13].

deflection of the T wave occurs. The normal time interval of a PR is 0.12 s to 0.2 s, QRS is 0.12 s and QT interval is up to 0.44 s. In the case of an atrial fibrillation, the ECG indicates an irregular rhythm pattern and a very high heart beat rate. The most significant indication of atrial fibrillation is the absence of P waves and PR intervals in the monitored ECG. On the other hand, Atrial Flutter (AFL) forms a unique pattern in ECG, which comprises saw tooth waves along with the absence of P waves [13].

P wave and QRS complex are depolarization waves in a normal ECG, whereas the T wave is a re-polarization wave [11]. Thus, the ECG signal is composed by segments which represent depolarization and re-polarization.

Heart depolarization begins in the sinoatrial node. However, in abnormal situations, this depolarization may begins in three other locals: atrial muscle, region around the atrioventricular node, or ventricular muscle. In the supraventricular rhythms, the depolarization wave scatters to the ventricles normally, and then the QRS complex presents its standard form. On the other hand, in the ventricular rhythms the depolarization wave scatters from the ventricles and then the QRS complex becomes large, and consequently presents a abnormal form. In this situation the T wave form is also abnormal [2].

In this work, seventeen different classes of ECG signals are evaluated. Fig. 2 highlights three examples from classes considered in this study: Atrial Fibrillation (AFIB), Atrial Flutter (AFL), and Ventricular Bigeminy (B). In order to improve the visibility of the three signals, they were shifted vertically, with the addition of distinct averages.

For the AFIB signal, the ECG indicates a rhythm pattern with a very high heart beat rate and the absence of P waves and PR interval, while for the AFL signal, the sawtooth waves are noted along with the absence of P waves [13].

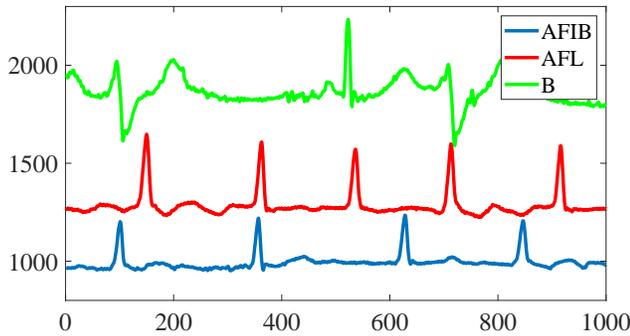


Figure 2. Examples of the cardiac rhythms: AFIB, AFL, and B.

III. DATABASE

This section covers the aspects related to the database used in this study, as well as the parameters applied in the ECG data acquisition.

The MIT-BIH Arrhythmia Database was used, which can be accessed in the public repository PhysioNet [10]. This database is composed by 48 records of two ECG channels, with duration of 30 *min.*, obtained from 47 people in the BIH Arrhythmia Laboratory (Beth Israel Hospital), between the years 1975 and 1979.

The data acquisition was performed with sampling frequency of 360 *samples/s* for each channel, with 11 bits per sample, which yielded a resolution of 10 *mV*. Two or more cardiologists verified the signals independently to assign annotations for the parts of the signals, in a total of 110,000 annotations, which were included in the database [15].

The table of rhythms in [16] summarizes the period of each rhythm in the acquisition process. A total of 45 records were selected because the records 102, 104, and 232 do not present the derivation modified limb lead II. As explained in section II, seventeen rhythms were selected to validate the proposed methodology.

IV. PREPROCESSING AND FEATURE EXTRACTION

This section presents the steps performed before the classification process, which comprises training, validation/optimization, and test steps. The first step concerns to the segmentation of ECG signals, because each record of the database contains more than one type of arrhythmia. In the second step, it is carried out the class balance. This step is necessary due to the great difference in the number of segments of each arrhythmia in the records. To minimize problems arising from data values ranging between different scales, it is necessary that multivariate data be standardized on the same scale, thus the next step is a statistical normalization. Finally, the last step of the preprocessing takes into account the feature extraction from the signals in order to feed the classification process.

Fig. 3 exhibits the steps involved in the whole methodology used in this work. This approach includes the following steps: (1) data acquisition; (2) data segmentation; (3) class balance; (4) statistical normalization; (5) feature extraction; (6) training and optimization; and (7) evaluation of results.

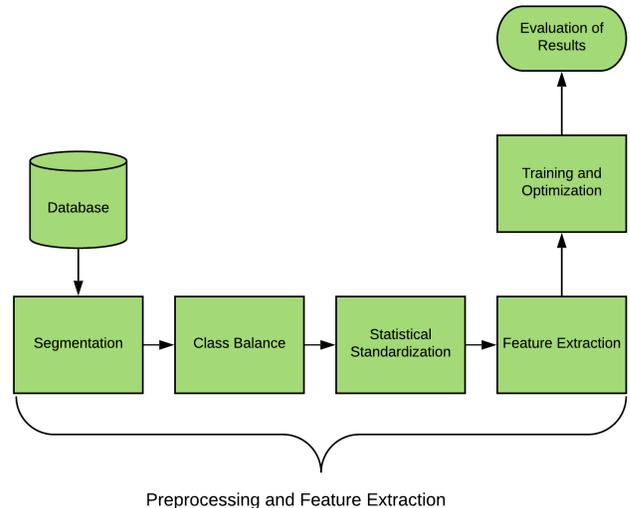


Figure 3. Block diagram of the proposed methodology.

A. Segmentation

The signals obtained from the PhysioNet repository can be composed by more than one rhythm. For example, the record 119 has the occurrence of two types of rhythms: Ventricular Bigeminy (B) and Ventricular Trigeminy (T). Thus, a data segmentation is necessary to separate fragments of this record.

In the segmentation step is performed the selection of segments with only one rhythm, which are 10 *s* long. Considering a sample rate of 360 *samples/s*, each segment is composed by 3600 samples.

For this task the toolbox PhysioNet Toolkit [17] was employed, which allows one to annotate each segment. By means of this annotation it is possible to identify the rhythm in each part of the signal. After segmentation, the data described in Table I were obtained.

Table I
DATA DESCRIPTION AFTER SEGMENTATION.

Id	Rhythm	#segments
1	Atrial Fibrillation (AFIB)	135
2	Atrial Flutter (AFL)	20
3	Atrial Premature Beat (APB)	66
4	Ventricular Bigeminy (B)	55
5	Fusion of Ventricular and Normal Beat (Fusion)	11
6	Idioventricular Rhythm (IVR)	10
7	Left Bundle Branch Block Beat (LBBBB)	103
8	Normal Sinus Rhythm (N)	283
9	Pacemaker Rhythm (PR)	45
10	Premature Ventricular Contraction (PVC)	133
11	Right Bundle Branch Block Beat (RBBBB)	62
12	Second-degree Heart Block (SDHB)	10
13	Supraventricular Tachyarrhythmia (SVTA)	13
14	Ventricular Trigeminy (T)	13
15	Ventricular Flutter (VFL)	10
16	Ventricular Tachycardia (VT)	10
17	Pre-excitation (WPW)	21

B. Class Balance

This step is carried out to make the number of segments per class as similar as possible, because high differences

in the number of segments among the classes might impair the performance of the classifiers [18], mainly in the optimization step in terms of the searching for optimal parameters. The algorithm employed to perform the balancing was ADASYN [19], which is an implementation of the Synthetic Minority Over-sampling Technique (SMOTE). ADASYN makes the new (synthetic) segments be located mainly in the board between the classes, and generates more synthetic data for the minority classes, because the learning might be compromised when performed with few segments. This algorithm includes the following steps:

- Evaluating of the degree of class imbalance, i.e., how much the majority class is greater than the minority ones;
- Calculating the number of synthetic segments to be created;
- Evaluating of the difficulty to learn each segment in minority classes, based on the proximity of these segments with segments from the majority class;
- Calculating the number of synthetic segments to be created for each original segment from the minority classes, based on the aforementioned metric, in which the segments with more segments in their neighborhood receive more synthetic segments. This step makes ADASYN to generate segments on the boundary of the classes, and to improve the overall learning;
- Generating of the synthetic segments such that

$$\hat{x}_i = x_i + (x_i - x_n)\lambda, \quad (1)$$

where x_i is the current segment from a minority class being evaluated, x_n is a randomly chosen segment of the minority class in the neighborhood of x_i , and λ is a random number such that $\lambda \in [0, 1]$.

ADASYN was designed to balance classes in binary problems, but the database considered in this study has seventeen classes. Then, an *one against all* strategy was employed to circumvent this issue. In this strategy, one class representing the minority one was chosen, and all the others composed the majority one. It was performed for all classes, with exception for the original majority class, Normal Sinus Rhythm (N).

C. Feature extraction

Frequency domain allows one to analyze many signal features, which in time domain might be difficult. One of these features is the Power Spectral Density (PSD). Through PSD one can infer how the signal energy is distributed in the frequency domain. This is specially crucial, as the PSD is essential to identify cardiac rhythm, as shown in [20], [21].

In this study, the method used to estimate the PSD was Welch [22]. In the time domain, this method divides the signal into successive blocks, forming a periodogram for each block, and then calculates the average of these periodograms. In other words, it is an average of periodograms over time. The advantages of the Welch's method include: the dimension reduction of the input of the classifiers, due to the final length of the spectrum to be defined by the window length; and the minor impact of the noise from the acquisition process, which is circumvented by the averaging of the segments.

The estimation of PSD via Welch's method was performed by considering three widths of Hamming's windows: 256, 512, and 1024 samples. Afterwards, the data were partitioned into three disjointed sets. This data split process was considered for each window width. The first set was used to train the classifiers, with 70% of total segments. Other data set comprised 15% of segments and was employed as a validation set to estimate the optimal parameters for each classifier. Finally, the last set with the remaining data was used to test the generalization performance of each classifier when new data were considered.

V. TRAINING, VALIDATION/OPTIMIZATION, AND TEST OF CLASSIFIERS

Four classifiers were chosen for this study. The first choice was a Neural Network (NN), the type of NN selected was the MLP. This technique is one of the most popular machine learning algorithms used in many different real-world applications. MLP is part of a general class of structures called feed-forward neural networks [23], and have been used in many problems involving modeling and optimization. In its structure, neurons are grouped into different layers. The first layer is called input, while the last layer is called output. The rest of the layers that lie between the input and output layers are called hidden layers. MLP can model complex non-linear functions in the training and validation steps, and generalize accurately new and unseen data in the test step.

The second choice was the SVM, another nonlinear algorithm by Cortes and Vapnik [24], [25]. The basic idea of the SVM's working principle is to find the hyperplane that can separate the data belonging to two classes with maximum margin; this is called the ideal hyperplane. The advantages of SVM over conventional classification methods are its greater generalization capability, its adaptability to various classification problems by changing kernel functions, and its optimal global solution.

The third choice involves an approach with working principle distinct from the previous two. The RF was proposed by Breiman based on the Decision Tree algorithm [26] and have been widely applied in the area of pattern recognition [27]. This technique is a machine learning method which is composed of many decision trees for classification and prediction. Features are randomly chosen for each decision tree to be trained, and the overall RF outputs are the result of voting by all decision trees.

The kNN is one of the well-known classification methods, becoming one of the ten most widely used data mining algorithms [28]. This algorithm requires the calculation of the distance of the unlabeled data for all labeled data in the training set. The kNN arose in 1951 from the need to perform a discriminant analysis when reliable parametric estimates of probability density were difficult to determine [29]. Over the years other kNN-based approaches have emerged [30]–[32].

As four different algorithms and three distinct widths of Hamming's windows were used, the total number of classifiers was twelve. In the validation step, the k-fold technique [33] was applied with ten folds, which generates ten distinct subsets

for training and test from the validation data set. This step was crucial to select the optimal parameter for each classifier. Afterwards, each classifier configured with its optimal parameter was trained on the training data set. Then, the trained classifier/model can be employed to classify new data from the test data set.

Note that, when the classes are balanced, one can expect that the validation step should be less complex in terms of computational effort to optimize a given parameter of a classifier, because there are more data examples to learn and extract relevant patterns.

VI. EVALUATION CRITERIA

To evaluate the classifiers, there are metrics often applied in many works [4], [34], [35]. In this paper, the following methods were chosen:

- Accuracy (AC): It determines how close the output of the classifier can be to the ground truth. The formulation is based on true-positive (TP), false-negative (FN), true-negative (TN), and false-positive (FP), such that:

$$AC = \left(\frac{TP + TN}{TP + FP + TN + FN} \right) \times 100\%. \quad (2)$$

- Sensitivity (SEN): It relates the segments correctly classified belonging to a class and the sum of these segments with the segments classified incorrectly belonging to another class, such that:

$$SEN = \left(\frac{TP}{TP + FN} \right) \times 100\%. \quad (3)$$

- Specificity (SPE): It presents the fraction of true-negatives, that is, the segments correctly classified as not belonging to a class in relation to the sum of these segments with the segments classified incorrectly belonging to another class, such that:

$$SPE = \left(\frac{TN}{TN + FP} \right) \times 100\%. \quad (4)$$

- Fleiss' kappa κ [19]: It is a coefficient used to evaluate the efficiency of a given classifier, its application is justified in problems of many classes. This coefficient is used when it is desired to discover the agreement between several classifiers, such that:

$$\kappa = \left(\frac{M \sum_{j=1}^L m_{j,j} - \sum_{j=1}^L G_j C_j}{M^2 - \sum_{j=1}^L G_j C_j} \right) \times 100\%, \quad (5)$$

where L is the number of classes; M is the total number of segments classified which are compared to the ground truth; $m_{j,j}$ is the number of segments belonging to the true class j that are also classified as class j (diagonal of the confusion matrix); C_j is the total number of segments classified belonging to the class j ; and G_j is the total number of ground truth segments belonging to the class j .

Table II
CLASSIFIERS' SETTINGS.

Classifier	Parameter description	Parameter values
SVM	Type	C-SVM
	Kernel Type	Radial base function
	γ (kernel bandwidth)	From 10^{-5} to 10^{-3} with step of 10^{-6}
kNN	Distance	Euclidean
	#neighbors	From 1 to 100
	Method	Classification
RF	Method	Classification
	#trees	From 1 to 300
MLP	Training algorithm	Scaled conjugate gradient backpropagation
	#neurons in the hidden layer	From 5 to 100 with step of 5
	#neurons in the output layer	17
	Input size	257, 513, and 1025

VII. RESULTS AND DISCUSSION

In this section, the obtained results are presented and discussed by applying the SVM, kNN, RF, and MLP, considering three widths of Hamming's window, on balanced and unbalanced databases composed of PSDs. The evaluation of the results was fair, according to section VI, by means of AC, SEN, SPE, and Fleiss' kappa κ employed on the test data set.

The results were generated using a notebook with Intel Core i5-8250U processor, 8GB RAM, MATLAB R2018a and LIBSVM library [36]. Scripts in MATLAB were developed for segmenting the ECG data, performing class balance by generating synthetic segments, normalizing the segments, extracting the PSDs, training, optimizing and testing the classifiers, and finally evaluating the results.

The techniques described in section V were configured as shown in Table II. Furthermore, the range of values considered for the parameters optimized in the validation step is also highlighted for each classifier.

A. Preprocessing Results

In relation to the class balance step, firstly the number of segments in class N (the original class with more segments) was fixed to 283 segments before the application of the ADASYN algorithm, in order to minimize the amount of synthetic segments created at the end of this step. Fig. 4 highlights the number of segments in each class along with the window width. One can note how the databases remained after the application of the class balance step. In general, there was an increase in the total number of segments in the databases, since as shown in Table I, the previous total number of segments was 1000. For each width of Hamming's window, the total number of segments were different; for windows of 256, 512, and 1024 samples, a total of 4617, 4689, and 4668 segments were reached. Besides, except class N that had no synthetic segment created, all other classes had the number of segments ranging between 241 and 282.

After class balance, the PSDs were extracted from the segments. Fig. 5 shows examples of PSDs for window of 1024 samples. The parameters used in the generation of these PSDs

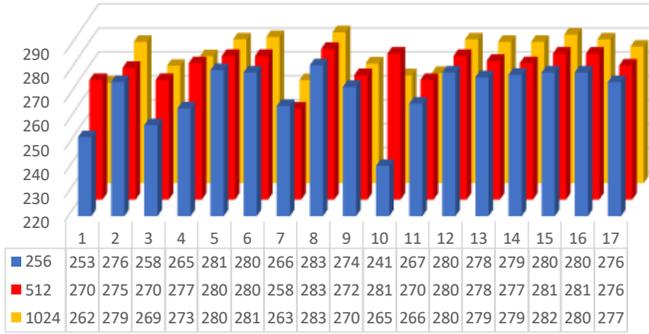


Figure 4. Number of segments in each class after class balance step. The blue, red, and yellow bars denote the segment obtained with the window widths of 256, 512, and 1024 samples, respectively, according to the Ids in Table I.

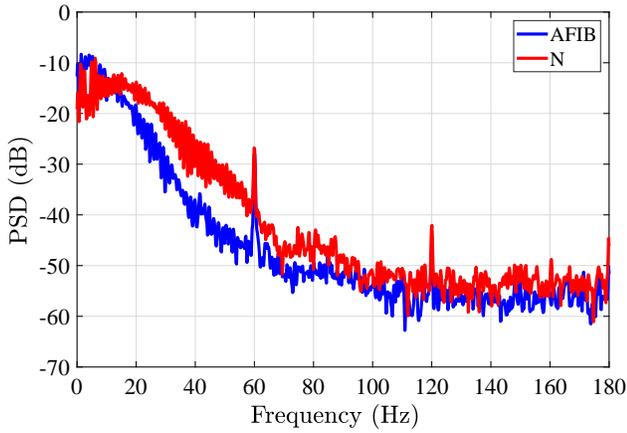


Figure 5. Examples of PSDs for classes AFIB and N, with the window width of 1024.

were: sampling frequency of 360 Hz and number of DFT points equal to 2048.

From Fig. 5, one can infer that the energy from the ECG signal is concentrated in low frequencies, which matches the literature observation. Thereby, 99.55% of the power is concentrated below 40 Hz, which can indicate the possibility of reducing the evaluated frequency band. It is also possible to identify two spikes at 60 Hz and 120 Hz, which probably represent the power line interference [37]. This interference might impair the time domain evaluation of the ECG, but in the PSD it acts like *bias* for the ML algorithms, and does not reduce the classification performance. One can note that the PSD from class AFIB has more energy in the frequencies of up to 10 Hz in comparison with the PSD from class N, whereas in all the rest of the frequency range, the former has less energy than the last. For instance, this can contribute to differentiate these two PSD examples through the classifiers.

B. Classification Results

Tables III, IV, V, VI, VII, and VIII reveal the classification results obtained with the application of the selected four machine learning algorithms in the test step, considering or not ADASYN for class balance. The most important results are

accuracy and sensitivity, as stated in [4], but another coefficient that is analyzed in this work is κ .

Table III
CLASSIFICATION RESULTS OBTAINED FOR WINDOW OF 256 SAMPLES WITH ADASYN.

Classifier	AC	SPE	SEN	κ
SVM	97.11%	99.82%	97.06%	73.90%
kNN	96.68%	99.79%	96.59%	69.98%
RF	97.54%	99.85%	97.54%	77.81%
MLP	98.15%	99.88%	98.23%	83.32%

Table IV
CLASSIFICATION RESULTS OBTAINED FOR WINDOW OF 256 SAMPLES WITHOUT ADASYN.

Classifier	AC	SPE	SEN	κ
SVM	80.66%	98.61%	71.49%	42.73%
kNN	82.66%	98.85%	78.06%	36.12%
RF	80.67%	98.64%	76.27%	42.73%
MLP	88.46%	99.11%	83.53%	14.21%

For the window of 256 samples, as shown in Table III, the results with better performance in the classification were obtained using the MLP, although the other classifiers achieved acceptable results in terms of AC, SPE, and SEN. From the optimization step, the parameter γ for the SVM classifier was 9.39×10^{-4} , for the kNN classifier the number of neighbors was 1, for the RF classifier the number of trees that generated the best results was 185, and for the MLP the number of neurons in the hidden layer was 95. On the other hand, one can note a clear decrease in performance from the results synthesized in Table IV, where there were no segments generated via ADASYN. With the absence of ADASYN, the metrics AC, SEN, and κ were mainly affected, being reduced approximately at least by 10%, 15%, and 35%, respectively.

The results for the windows of 512 samples are shown in Table V. The increase of the window width to 512 samples showed that, in a general perspective, this increase improved the classification performance for all algorithms. For this window width, the RF classifier performed as the best algorithm, with the number of trees as 134. The results from MLP were marginally greater than those obtained with the former window width and competitive with those from the RF. One can note that the kNN classifier had the worse performance, being the optimal number of neighbors equal to 1. The SVM classifier showed an improvement with the increase of the number of samples used in this window, but its results in terms of AC, SEN, and κ were relatively far from the best ones. The parameter γ needed to achieve this performance was 6.21×10^{-4} . For the MLP classifier, the number of neurons in the hidden layer was 70 after the optimization step. Again, a poor performance can be highlighted in Table VI, without the application of ADASYN. Despite some improvements in AC and SPE, the drawbacks yielded from the unbalance of the original data emerged when SEN and κ were estimated.

The classification results obtained for the window of 1024 samples are highlighted in Table VII. The best results were obtained by employing the SVM classifier, with γ as 1.6×10^{-4} .

Table V
CLASSIFICATION RESULTS OBTAINED FOR WINDOW OF 512 SAMPLES WITH ADASYN.

Classifier	AC	SPE	SEN	κ
SVM	97.72%	99.86%	97.74%	79.44%
kNN	97.29%	99.83%	97.30%	75.59%
RF	98.72%	99.92%	98.73%	88.44%
MLP	98.42%	99.90%	98.35%	85.74%

Table VI
CLASSIFICATION RESULTS OBTAINED FOR WINDOW OF 512 SAMPLES WITHOUT ADASYN.

Classifier	AC	SPE	SEN	κ
SVM	85.33%	98.95%	62.77%	24.50%
kNN	87.33%	99.14%	85.76%	12.58%
RF	87.33%	99.06%	80.68%	12.58%
MLP	89.53%	99.19%	82.32%	31.21%

Table VII
CLASSIFICATION RESULTS OBTAINED FOR WINDOW OF 1024 SAMPLES WITH ADASYN.

Classifier	AC	SPE	SEN	κ
SVM	98.86%	99.93%	98.85%	89.68%
kNN	96.00%	99.75%	96.08%	63.88%
RF	97.14%	99.82%	97.18%	74.20%
MLP	98.31%	99.90%	98.31%	84.78%

Table VIII
CLASSIFICATION RESULTS OBTAINED FOR WINDOW OF 1024 SAMPLES WITHOUT ADASYN.

Classifier	AC	SPE	SEN	κ
SVM	84.66%	98.83%	67.88%	27.78%
kNN	86.67%	99.07%	86.52%	16.95%
RF	83.33%	98.68%	76.61%	33.56%
MLP	90.26%	99.17%	83.15%	38.06%

This window width yielded the best performance for the SVM, which can be explained by the improvement in the frequency resolution [38]: $\Delta f_{1024} \approx 0.175 \text{ Hz/bin}$, against $\Delta f_{512} \approx 0.351 \text{ Hz/bin}$ and $\Delta f_{256} \approx 0.703 \text{ Hz/bin}$. With the improvement in the frequency resolution, more discriminant the extracted feature can be, which tends to aid the SVM classifier in separating the segments per classes correctly, at the cost in the increase of the input dimension for this algorithm. The MLP obtained average results for the window of 1024 samples in comparison with the other two window widths, the number of neurons in the hidden layer was 55 after the optimization step. For the kNN and RF, the best results were obtained with the number of neighbors and number of trees equal to 1 and 208, respectively, and in this manner it is possible to observe that an increase of the window width negatively influences their performances. Without the application of ADASYN, Table VIII provides the better results reached in terms of AC and SEN, which were yield by the MLP and kNN, respectively. However, even though these results were better, they still fall behind the results supported by ADASYN for this window, with differences approximately ranging between 8% and 12%.

The reason related to the general lower performance of kNN would be the fact this is not a good classifier for data with nonlinearities. On the other hand, the superior performance reached by the other algorithms, with distinct window widths, can be explained by the fact that their capabilities do not decline for high-dimensional data, as demonstrated in [39]. With the exception of kNN, the other classifiers can perform very well for nonlinear data, the RF's performance was lower than MLP and SVM in some cases because its classification approach is more suitable for categorical data and the problem addressed in this work did not favor this characteristic. A possible explanation for the difference in performance between MLP and SVM is the fact that MLP has many parameters which might be optimized and in our approach only the number of neurons in the hidden layer was optimized, which may have contributed to its performance being marginally lower than SVM for the window of 1024 samples.

A comparison between the results of this paper and other

works from the scientific literature is carried out in Table IX. It is clear that the SVM has been the most employed classifier for arrhythmia recognition. According to the results of this work and [4], the SVM and MLP proved to be superior to the other techniques tested. It is also evident that the application of ADASYN can improve the classification performance of unbalance problems, even though this technique generates more data to be classified. Nevertheless, in relation to the computational cost, the work [4] adjusted the parameters of the classifiers with a genetic algorithm, which caused the consumption of several hours in the optimization step, whereas our approach consumed only few minutes per classifier. This fact elucidates that the proposed methodology has the potential to achieve a satisfactory performance with less computational resources. It is also important to emphasize that all works in Table IX used the data available in the Physionet repository [10], which indicates that the signals treated in these works were acquired in the same conditions.

VIII. CONCLUSIONS

In this paper, a novel methodology based on digital signal processing and machine learning was proposed for automatically classifying different cardiac rhythms. This approach comprised the following steps: (1) segmentation, (2) class balance with ADASYN, (3) statistical normalization, (4) feature extraction, (5) training, optimizing, and testing of classifiers, and (6) evaluation of results. The main characteristic of this methodology was the application of ADASYN algorithm to balance the number of segments between the classes in order to simplify the parameter optimization of kNN, SVM, RF, and MLP. This procedure impacted positively the classification performance as the addition of more segments in a minority class became it more representative in terms of a distinct pattern. Besides, this methodology also demonstrated promising results when compared with the state-of-the-art works.

As future works, the proposed methodology will be applied in databases with more classes, different types of features will be extracted with the aim of minimizing the input size of the classifiers, and nonparametric statistical tests will be conducted to attest the significance in the classification performance.

Table IX
A COMPARISON WITH RELATED WORKS.

Authors	Classifiers	#classes	AC	SPE	SEN	κ
Butt, Akram and Khan [3]	SVM	15	98.78%	-	-	-
Rezaei <i>et al.</i> [5]	kNN	4	97.31%	98.21%	94.64%	-
Plawiak [4]	SVM, kNN, Probabilistic Neural Network and Radial Basis Function Neural Network	13, 15 and 17	98.85% for 17 classes	99.39% for 17 classes	90.20% for 17 classes	88.49% for 17 classes
Pinho, Gomes and Santos	SVM, kNN, Random Forest and MLP	17	98.86% for 1024 samples	99.93% for 1024 samples	98.85% for 1024 samples	89.68% for 1024 samples

REFERENCES

- [1] World Heart Organization, "World heart day," last accessed 19 September 2019. [Online]. Available: https://www.who.int/cardiovascular_diseases/world-heart-day/en/.
- [2] J. Hampton, *The ECG made easy*, 8th ed. Theobalds Road, London, UK: Churchill Livingstone, 2013.
- [3] M. M. Butt, M. U. Akram, and S. A. Khan, "Classifying normal sinus rhythm and cardiac arrhythmias in ecg signals using statistical features in temporal domain," in *2015 9th Asia Modelling Symposium (AMS)*, Sep. 2015, pp. 28–31. [Online]. Available: <https://doi.org/10.1109/AMS.2015.14>
- [4] P. Plawiak, "Novel methodology of cardiac health recognition based on ECG signals and evolutionary-neural system," *Expert Systems with Applications*, vol. 92, pp. 334 – 349, 2018. [Online]. Available: <https://doi.org/10.1016/j.eswa.2017.09.022>
- [5] S. Rezaei, S. Moharreri, M. Abdollahpur, and S. Parvaneh, "Heart arrhythmia classification using extracted features in poincare plot of RR intervals," in *2017 Computing in Cardiology (CinC)*, Sep. 2017, pp. 1–4. [Online]. Available: <https://doi.org/10.22489/CinC.2017.115-399>
- [6] S. RUSSELL and P. NORVIG, *Artificial Intelligence: A Modern Approach*, 4th ed. Upper Saddle River, NJ, USA: Prentice Hall, 2009.
- [7] M. A. Escalona-Morán, M. C. Soriano, I. Fischer, and C. R. Mirasso, "Electrocardiogram classification using reservoir computing with logistic regression," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 3, pp. 892–898, May 2015. [Online]. Available: <https://doi.org/10.1109/JBHI.2014.2332001>
- [8] Y. Bazi, N. Alajlan, H. AlHichri, and S. Malek, "Domain adaptation methods for ECG classification," *2013 International Conference on Computer Medical Applications (ICCM)*, pp. 1–4, Jan 2013. [Online]. Available: <https://doi.org/10.1109/ICCM.2013.6506156>
- [9] A. K. Das, F. Catthor, and S. Schaafsma, "Heartbeat classification in wearables using multi-layer perceptron and time-frequency joint distribution of ECG," *IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*, pp. 69–74, Sep. 2018. [Online]. Available: <https://doi.org/10.1145/3278576.3278598>
- [10] M. L. for Computational Physiology, "Physionet: The research resource for complex physiologic signals." [Online]. Available: <https://physionet.org/>
- [11] J. E. Hall, *Guyton and Hall Textbook of Medical Physiology*, 13th ed. Philadelphia, PA, USA: Saunders, 2015.
- [12] A. M. Al-Busaidi and L. Khriji, "Digitally filtered ECG signal using low-cost microcontroller," *International Conference on Control, Decision and Information Technologies (CoDIT)*, pp. 258–263, May 2013. [Online]. Available: <https://doi.org/10.1109/CoDIT.2013.6689554>
- [13] P. Ganesan, "Characterization of cardiac electrogram signals during atrial fibrillation," Ph.D. dissertation, Rochester Institute of Technology, June 2015.
- [14] B. Surawicz and T. Knilans, *Chou's electrocardiography in clinical practice: adult and pediatric*, 6th ed. Amsterdam, NL: Elsevier Health Sciences, 2008.
- [15] G. B. Moody and R. G. Mark, "MIT-BIH arrhythmia database," last accessed 19 September 2019. [Online]. Available: <https://physionet.org/physiobank/database/mitdb/>.
- [16] —, "Tables of beats and rhythms," last accessed 19 September 2019. [Online]. Available: <https://physionet.org/physiobank/database/html/mitdbdir/tables.htm#allrhythms>.
- [17] —, last accessed 19 September 2019. [Online]. Available: <https://www.physionet.org/physiotools/matlab/wfdb-app-matlab/>.
- [18] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "Smote: synthetic minority over-sampling technique," *Journal of artificial intelligence research*, vol. 16, pp. 321–357, 2002. [Online]. Available: <https://doi.org/10.1613/jair.953>
- [19] H. He, Y. Bai, E. A. Garcia, and S. Li, "ADASYN: Adaptive synthetic sampling approach for imbalanced learning," *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, pp. 1322–1328, June 2008. [Online]. Available: <https://doi.org/10.1109/IJCNN.2008.4633969>
- [20] J. Surda, S. Lovas, J. Pucik, and M. Jus, "Spectral properties of ECG signal," in *2007 17th International Conference Radioelektronika*, April 2007, pp. 1–5. [Online]. Available: <https://doi.org/10.1109/RADIOELEK.2007.371653>
- [21] S. Das and M. Chakraborty, "Comparison of power spectral density (PSD) of normal and abnormal ECGs," *IJCA, Special Issue on 2nd National Conference-Computing, Communication and Sensor Network*, no. 2, pp. 10–14, 2011.
- [22] P. Welch, "The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," *IEEE Transactions on Audio and Electroacoustics*, vol. 15, no. 2, pp. 70–73, June 1967. [Online]. Available: <https://doi.org/10.1109/TAU.1967.1161901>
- [23] F. Scarselli and A. C. Tsoi, "Universal approximation using feedforward neural networks: A survey of some existing methods, and some new results," *Neural networks*, vol. 11, no. 1, pp. 15–37, 1998. [Online]. Available: [https://doi.org/10.1016/S0893-6080\(97\)00097-X](https://doi.org/10.1016/S0893-6080(97)00097-X)
- [24] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, Sep. 1995. [Online]. Available: <https://doi.org/10.1007/BF00994018>
- [25] V. Vapnik, *The nature of statistical learning theory*, 2nd ed. Midtown Manhattan, NY, USA: Springer-Verlag New York, 2000. [Online]. Available: <https://doi.org/10.1007/978-1-4757-3264-1>
- [26] L. Breiman, *Classification and Regression Trees*. New York, NY, USA: Routledge, 1984. [Online]. Available: <https://doi.org/10.1201/9781315139470>
- [27] —, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, Oct. 2001. [Online]. Available: <https://doi.org/10.1023/A:1010933404324>
- [28] N. S. Altman, "An introduction to kernel and nearest-neighbor nonparametric regression," *The American Statistician*, vol. 46, no. 3, pp. 175–185, 1992. [Online]. Available: <https://doi.org/10.1080/00031305.1992.10475879>
- [29] E. Fix and J. L. Hodges Jr, "Discriminatory analysis-nonparametric discrimination: consistency properties," California Univ Berkeley, Tech. Rep., 1951.
- [30] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Transactions on Information Theory*, vol. 13, no. 1, pp. 21–27, January 1967. [Online]. Available: <https://doi.org/10.1109/TIT.1967.1053964>
- [31] S. A. Dudani, "The distance-weighted k-nearest-neighbor rule," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-6, no. 4, pp. 325–327, April 1976. [Online]. Available: <https://doi.org/10.1109/TSMC.1976.5408784>
- [32] J. M. Keller, M. R. Gray, and J. A. Givens, "A fuzzy k-nearest neighbor algorithm," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-15, no. 4, pp. 580–585, July 1985. [Online]. Available: <https://doi.org/10.1109/TSMC.1985.6313426>
- [33] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, vol. 2, pp. 1137–1143, 1995. [Online]. Available: <https://dl.acm.org/doi/10.5555/1643031.1643047>

- [34] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861 – 874, 2006, ROC Analysis in Pattern Recognition. [Online]. Available: <https://doi.org/10.1016/j.patrec.2005.10.010>
- [35] J. L. Fleiss, "Measuring nominal scale agreement among many raters," *Psychological bulletin*, vol. 76, no. 5, pp. 378–382, 1971. [Online]. Available: <https://doi.org/10.1037/h0031619>
- [36] C.-C. Chang and C.-J. Lin. LIBSVM – a library for support vector machine. [Online]. Available: <https://www.csie.ntu.edu.tw/~cjlin/libsvm/>
- [37] L. Sörnmo and P. Laguna, *Bioelectrical signal processing in cardiac and neurological applications*. Burlington, MA, USA: Academic Press, 2005. [Online]. Available: <https://doi.org/10.1016/B978-0-12-437552-9.X5000-4>
- [38] S. V. Vaseghi, *Advanced Digital Signal Processing and Noise Reduction*, 3rd ed. Hoboken, NJ, USA: Wiley, 2006. [Online]. Available: <https://doi.org/10.1002/9780470740156>
- [39] I. Guyon, S. Gunn, M. Nikravesh, and L. A. Zadeh, *Feature extraction: foundations and applications*. New York, NY, USA: Springer, 2008. [Online]. Available: <https://doi.org/10.1007/978-3-540-35488-8>



N. Pinho received the B.Sc. degree in computer engineering from the Universidade Federal do Sul e Sudeste do Pará, Marabá, Brazil, in 2019.

His current research interests include artificial intelligence, programming, intelligent systems, telecommunications, and computer networks.



D. Azevedo received the M.Sc. and Ph.D. degrees in electrical engineering (telecommunications) from Universidade Federal Pará, Brazil, in 2012 and 2017, respectively. He is currently a Professor with the Faculty of Computing and Electrical Engineering, Institute of Geosciences and Engineering, Universidade Federal do Sul e Sudeste do Pará (Unifesspa).

His current research interests include MIMO communication systems and digital signal processing applied to machine learning.



A. Santos received the M.Sc. and Ph.D. degrees in electrical engineering (telecommunications and applied computing) from Universidade Federal do Pará, Belém, Brazil, in 2014 and 2017, respectively. He is currently a Professor with the Faculty of Computing and Electrical Engineering, Institute of Geosciences and Engineering, Universidade Federal do Sul e Sudeste do Pará (Unifesspa).

His current research interests include machine learning and structural health monitoring.