



# A Fast Machine Learning Model for ECG-Based Heartbeat Classification and Arrhythmia Detection

Miquel Alfaras\*, Miguel C. Soriano and Silvia Ortin

Instituto de Física Interdisciplinar y Sistemas Complejos, IFISC (UIB-CSIC), Palma de Mallorca, Spain

## OPEN ACCESS

### Edited by:

Raul Vicente,  
Max-Planck-Institut für Hirnforschung,  
Germany

### Reviewed by:

Haroldo Valentin Ribeiro,  
State University of Maringá, Brazil  
Reinaldo Roberto Rosa,  
National Institute of Space Research  
(INPE), Brazil

### \*Correspondence:

Miquel Alfaras  
m.phy@live.com

### Specialty section:

This article was submitted to  
Interdisciplinary Physics,  
a section of the journal  
Frontiers in Physics

**Received:** 14 May 2019

**Accepted:** 03 July 2019

**Published:** 18 July 2019

### Citation:

Alfaras M, Soriano MC and Ortin S  
(2019) A Fast Machine Learning  
Model for ECG-Based Heartbeat  
Classification and Arrhythmia  
Detection. *Front. Phys.* 7:103.  
doi: 10.3389/fphy.2019.00103

We present a fully automatic and fast ECG arrhythmia classifier based on a simple brain-inspired machine learning approach known as Echo State Networks. Our classifier has a low-demanding feature processing that only requires a single ECG lead. Its training and validation follows an inter-patient procedure. Our approach is compatible with an online classification that aligns well with recent advances in health-monitoring wireless devices and wearables. The use of a combination of ensembles allows us to exploit parallelism to train the classifier with remarkable speeds. The heartbeat classifier is evaluated over two ECG databases, the MIT-BIH AR and the AHA. In the MIT-BIH AR database, our classification approach provides a sensitivity of 92.7% and positive predictive value of 86.1% for the ventricular ectopic beats, using the single lead II, and a sensitivity of 95.7% and positive predictive value of 75.1% when using the lead V1'. These results are comparable with the state of the art in fully automatic ECG classifiers and even outperform other ECG classifiers that follow more complex feature-selection approaches.

**Keywords:** Echo State Networks, reservoir computing, arrhythmia classification, GPU, ECG

## 1. INTRODUCTION

Electrocardiogram (ECG) analysis has been established at the core of cardiovascular pathology diagnosis since its development in the twentieth century. The ECG signals reflect the electrical activity of the heart. Thus, heart rhythm disorders or alterations in the ECG waveform are evidences of underlying cardiovascular problems, such as arrhythmias. Non-invasive arrhythmia diagnosis is based on the standard 12-lead electrocardiogram, which measures electric potentials from 10 electrodes placed at different parts of the body surface, six in the chest and four in the limbs. In order to provide an effective treatment for arrhythmias, an early diagnosis is important. Early detection of certain types of transient, short-term or infrequent arrhythmias requires long-term monitoring (more than 24 h) of the electrical activity of the heart. The fast development of the digital industry has allowed for improvements in devices, data acquisition and computer-aided diagnosis methods.

The open access to ECG databases [1] has led to the development of many methods and approaches for computer-aided ECG arrhythmia classification over the last decades, fostering the productive cross-disciplinary efforts that engineers, physicists or non-linear dynamics researchers are no strangers to. Almost every computer-aided ECG classification approach involves four main steps, namely, the preprocessing of the ECG signal, the heartbeat detection, the feature extraction and selection and finally the classifier construction. The preprocessing of the ECG signal and the heartbeat detection are out of the scope of this work, both widely studied, and the heartbeat detection is close to optimal results [2].

A large number of classifiers have been proposed for arrhythmia discrimination. The proposed techniques range from simple classifiers, such as linear discriminants (LD) [3–5] or decision trees [5–7], to more sophisticated ones, such as traditional neural networks [8–13], Support Vector Machines (SVM) [9, 14–18], conditional random fields [19], and more recently deep learning techniques [13, 20–22]. In addition, many works have been devoted to finding the best combination of features, sometimes even developing complex signal processing methods, and to choosing the best subset (dimensionality reduction) for the arrhythmia classification [23]. On the one hand, popular choices for the input features are morphological features extracted from the time domain (such as inter-beat intervals, amplitudes, areas) [3, 14, 15, 24], frequency-domain features [6, 7, 16, 17, 25], wavelet transforms [4, 8–11, 18, 26], complex heartbeat representations [16] or higher order statistics (HOS) [4, 6, 7, 9]. On the other hand, feature selection methods, such as the independent component analysis (ICA) [18, 26], principal component analysis (PCA) [18], particle swarm optimization (PSO) [16], or the genetic algorithm—back propagation neural networks (GA-BPNN) [23], have been used.

Despite the good performance in classifying arrhythmias achieved by these methods, many of them require long computation times to optimize the classifiers. The use of complex classification or preprocessing methods is not suitable for online calculations or demand a lot of computational power. In this work, we present a fully automatic and fast classifier of arrhythmias that can be implemented online and analyze long sequences of ECG records efficiently. By loosening the requirements for feature extraction, we propose an implementation fundamentally based on raw signals, single lead information and heart rates that aims at reducing computation time while achieving low error classification results.

Cardiologists use mostly the raw ECG to diagnose. The simplest and fastest method of feature extraction is then to extract sampled points from an ECG signal curve. However, one should be aware of the fact that the amount of the extracted features used to characterize the heartbeat can be a burden for the classification algorithm. For this reason, most of the works that use the raw signal perform a down sampling of the waveform or some feature selection in order to reduce the computation time [3, 4, 15]. In order to circumvent this issue, a simple machine learning method is chosen to classify the arrhythmias. One of the advantages of the proposed method is that the number of features barely affects the speed of the classification since the classifier parameters related to the input are not optimized and remain random, as it will be described in more detail later in the text. As a result, the raw waveform of the heartbeat can be used for the classification without compromising speed. This simple machine learning method also allows a fast retraining of the classifier if new ECG data become available.

In this work, we propose an ensemble of Echo State Networks (ESNs) [27] as the classifier method, using the raw ECG waveforms and time intervals between the heartbeats as the input features. A particular advantage of the ESNs is that they

have recurrent connections, being able to take into account time dependencies between neighboring heartbeats. This property is beneficial since, in the case of a normal or an abnormal heartbeat, there are more chances that the subsequent heartbeat will also be a healthy or a pathological one. Moreover, the ESN method can take advantage of the power of a parallel computing architecture, such as a graphics processing unit (GPU). Hence, we compare the computation times between a GPU and a central processing unit (CPU), showing that the implementation in a GPU outperforms its CPU counterpart in the classification of the heartbeats. The computation times of the GPU outperform those of the CPU even in the training part of the classifier, i.e., the entire system can be trained extremely fast with a GPU.

Finally, it is worth noting that our classifier is based on a single lead ECG. Long-term monitoring generally involves devices with fewer electrodes than the standard 12 leads ECG in order to allow the patient to have a normal activity, requiring computer-aided techniques to analyze the huge amounts of data generated. We show that our heartbeat classification method outperforms other classifiers that rely on much more complicated feature selection techniques and complex calculations. We evaluate the proposed classifier in two different ECG databases and leads to test the robustness of the proposed algorithm.

## 2. MATERIALS AND METHODS

### 2.1. Databases

The performance of the proposed heartbeat classification method has been evaluated in two internationally recognized ECG databases: the MIT-BIH arrhythmia (MIT-BIH AR) [28] and the AHA [29]. The MIT-BIH AR database is a golden standard to evaluate arrhythmia classifiers. This benchmark database consists of 48 half-hour ECG records sampled at 360 Hz. Each ECG record contains two leads: lead II (modified limb lead II, obtained from electrodes on the chest) and lead V1' (modified lead V1, and in some records V2, V4, or V5). The AHA database contains 154 ECG recordings of 3 h long but only the last 30 min have information about the beat class. The AHA ECG recordings have two leads (A,B) sampled at 250 Hz. The documentation of the AHA database does not provide the name of the leads.

Both databases have annotations indicating the class of the heartbeat and its position verified by independent experts. Following the standards and recommendations of the American National Standards Institute developed by the Association for the Advancement of Medical Instrumentation (AAMI) for the evaluation of ECG classifiers [30], all the heartbeat annotation labels are converted to five heartbeat types: N (normal beats), S (supraventricular ectopic beats), V (ventricular ectopic beats), F (fusion beats), and Q (unclassifiable beats). The Q beats were excluded in this research because they are not representative [31]. Also in accordance to the AAMI standard, ECG recordings with paced beats are removed (i.e., four ECG records in the MIT-BIH AR database and three ECG records in the AHA database are excluded from the analysis). It is worth mentioning that the original annotations of the AHA database do not differentiate between N and S beats.

### 2.1.1. Training and Test Datasets

Each database is split into two sets: one for training (DS1) and one for testing (DS2). This division of the data is chosen to balance the presence of the different types of heartbeats and number of subjects in each dataset. It takes into account the inter-patient division, i.e., the subjects used to construct or optimize the classifier (DS1) are different from the subjects used to evaluate it (DS2). It has been demonstrated [3] that models which use heartbeats of the same patient in both the training and test are biased and their results can not be replicated in real environments.

For the MIT-BIH AR database we adopted the same set division as in de Chazal et al. [3] for comparative purposes of the results. 22 of the 44 ECG records of the MIT-BIH AR database are part of the set DS1 and the other 22 are part of the set DS2. For the AHA database, we use the recordings recommended for the training and testing procedure in the original AHA database description. In the AHA database, the set DS1 contains 79 ECG recordings with the label *series* = 0 and the DS2, 75 recordings labeled with *series* = 1. The division scheme for the MIT-BIH AR and AHA databases is summarized in **Tables 1, 2**, respectively. The beat class distributions of the different databases are given in **Table 3**.

## 2.2. Performance Metrics

The performance of the proposed algorithm is evaluated using the MIT-BIH AR and AHA databases on a single lead basis. The performance of each classification algorithm is assessed using four standard statistical measures: sensitivity (*Se*), positive predictive value (PPV), specificity (*Sp*), and accuracy (*Acc*). They are calculated as follows:

$$Se = TP/(TP + FN), \quad (1)$$

$$PPV = TP/(TP + FP), \quad (2)$$

$$Sp = TN/(TN + FP), \quad (3)$$

$$Acc = (TP + TN)/(TP + TN + FP + FN) \quad (4)$$

True positives (TP) indicate correctly predicted positive class and true negatives (TN) indicate correctly predicted negative class heartbeats. A good classifier is the one that minimizes false negatives (FN) and false positives (FP).

The F1 score is the harmonic mean of *Se* and PPV,  $F1 = 2(Se \cdot PPV)/(Se + PPV)$ . The F1 score is used to choose the optimum parameters of our classifier during the training phase.

## 2.3. The Heartbeat Classifier

The proposed heartbeat classifier is based on an Echo State Network (ESN). It classifies the heartbeats of the processed ECG recordings in two classes based on morphology: SVEB+ and VEB+. SVEB+ class includes normal (N) and supraventricular ectopic (S or SVEB) heartbeats. These heartbeats have a normal morphology and a supraventricular origin as opposed to VEB+ heartbeats that present ventricular origin or abnormal

**TABLE 1** | Distribution of the MIT-BIH AR database ECG recordings into the training (DS1) and testing (DS2) sets.

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

**TABLE 2** | Distribution of the AHA database ECG recordings into the training (DS1) and testing (DS2) sets.

Dataset	AHA recordings
DS1	1,001–1,004, 1,006–1,010, 2,001, 2,003–2,010, 3,001–3,010, 4,001–4,010, 5,001–5,010, 6,001–6,010, 7,001–7,010, 8,001–8,004, 8,006–8,010
DS2	1,101–1,110, 2,101–2,110, 3,101–3,110, 4,101–4,110, 5,101–5,105, 6,101–6,110, 7,101–7,110, 8,101–8,105, 8,107–8,110

The ECG recording names in the AHA database are of the form CSNN, where C is the arrhythmia category, S is the series and NN is the file number in the category.

**TABLE 3** | Heartbeat class distribution of the training (DS1) and testing (DS2) sets.

Database	SVEB+ class		VEB+ class	
	N	S	V	F
MIT-BIH AR (DS1)	45,783	943	3,785	414
MIT-BIH AR (DS2)	44,179	1,834	3,216	388
AHA (DS1)	158,587		15,075	292
AHA (DS2)	156,992		15,855	437

The beats at the beginning and at the end of the recordings are discarded as they do not provide information about the temporal distance to the neighboring beats.

morphology. The VEB+ class comprises the ventricular ectopic beats (V or VEB) and the fusion beats (F).

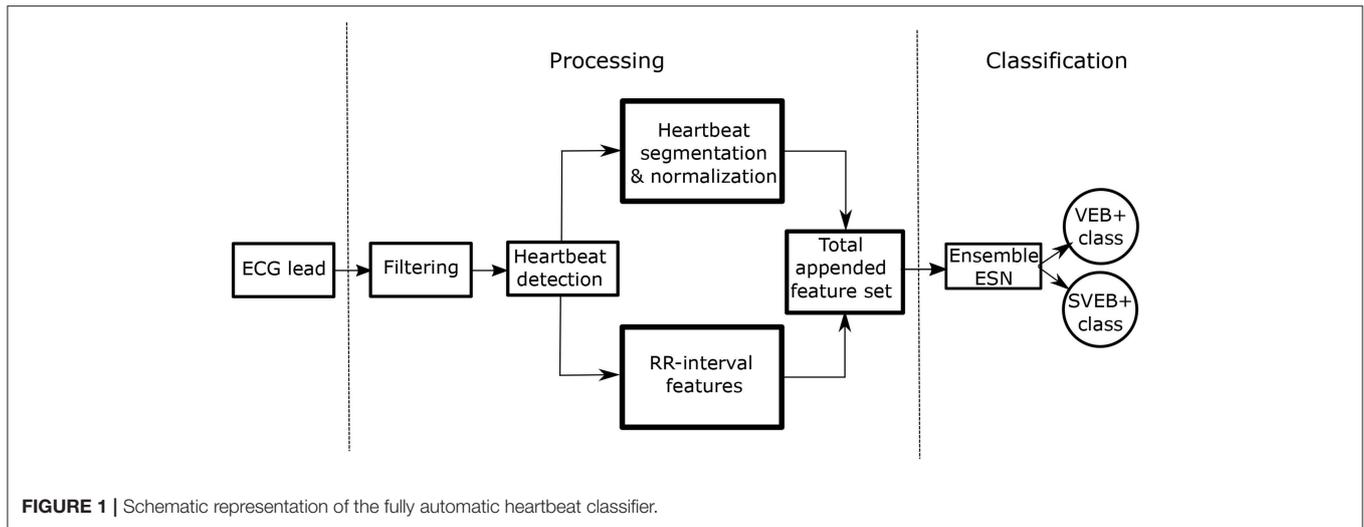
The overall process is schematically represented in **Figure 1**. The two stages are clearly differentiated:

- Stage 1–Processing of the ECG recordings: this procedure involves the filtering, heartbeat detection, heartbeat segmentation, and feature extraction. We include morphological and time intervals between heartbeats in our model.
- Stage 2–Classification between SVEB+ and VEB+ classes: we use an ensemble of ESNs with ring topology to perform this classification task.

We discuss the classification procedure in stage 2 in more detail later in the text.

## 2.4. Processing of the ECG and Feature Extraction

In order to accomplish arrhythmia classification, minor preprocessing needs to be applied to the source ECG records. In



our system, the processing of the ECG recordings includes the following steps:

1. **ECG re-sampling:** ECG signals are processed with a common sampling rate of 250 Hz. The AHA database (250 Hz) keeps its original sampling rate and the MIT-BIH AR database (360 Hz) is resampled to 250 Hz using the PhysioToolkit software package [1].
2. **ECG filtering:** All ECG recordings are filtered in a bandwidth  $\nu$  (Hz)  $\in [0.5, 35]$ , to correct the baseline and remove unwanted high frequency noise. A Butterworth high-pass filter (with a cutoff frequency  $\nu_c = 0.5$  Hz) and a finite impulse response filter of 12th order (35 Hz, at 3-dB point) are used, following standard procedure.
3. **Heartbeat detection:** To determine the position of the heartbeats, the annotated positions provided by the databases are used. In the MIT-BIH AR database the annotation position occurs at the largest of the local extrema of the QRS complex. Beat detection is beyond the scope of this study. Highly accurate automated beat detection methods have already been reported [32].
4. **RR calculation:** The RR interval is defined as the time interval between successive heartbeats. The RR interval associated to a heartbeat  $i$ ,  $RR(i)$ , corresponds to the time difference between the heartbeat  $i$  and the previous heartbeat ( $i - 1$ ).
5. **Heartbeat segmentation:** The ECG signal is segmented around the annotated position given by each database. The size of the segmented heartbeat is 240 ms (60 samples at 250 Hz) and it is centered around the annotation position.
6. **Heartbeat normalization:** Each segmented heartbeat is normalized between  $[-1, 1]$ . This scaling operation results in a signal that is independent of the original ECG recording amplitude.

After processing the ECG recordings, each heartbeat is represented by a set of features. One of the main goals related to the feature selection in our model is to avoid complicated features with a high computational cost, since we aim to design a fast and real-time heartbeat classifier. Therefore, we focus on simple ways

to extract features. In our case, we use the raw waveform of each heartbeat around the heartbeat position to represent it. The raw data of each beat was represented by an equal number of samples from each side from the point of the beat annotation. In order to learn from the temporal characteristics of each beat, information about the RR intervals is also added to the heartbeat features. The RR intervals are features used in almost all the methods to classify arrhythmic heartbeats. For instance, it is well-known that VEB heartbeats are characterized by shorter RR intervals than the N heartbeats. We found that using the logarithm of the RR intervals, as in Llamedo and Martinez [33], leads to a slightly better performance of the classifier. All the features that characterize the  $i$ th heartbeat are listed below:

- 60 raw samples of the segmented heartbeat waveform centered around the position annotated for the heartbeat.
- $\ln(RR(i))$ : logarithm of the current RR interval.
- $\ln(RR(i + 1))$ : logarithm of the next RR interval.
- $\ln(RR_{mean})$  logarithm of an average of the previous 250 RR intervals (averaging over the  $n$  available RR intervals when  $n < 250$ ).

At the end of the processing and feature extraction stage, each heartbeat is represented as a  $d$ -dimensional vector containing three features related to the RR intervals and 60 morphological features, which are simply the samples of the ECG waveform around the position annotated for each heartbeat. This  $d$ -dimensional vector ( $d = 63$ ) is the input for the classification algorithm.

## 2.5. Classification Algorithm: Echo State Network

Our classifier is built upon an ESN with a ring topology. ESNs are a popular implementation of Reservoir Computing (RC). RC is an established paradigm in machine learning that has been successfully applied in a variety of different tasks [27, 34]. This computing paradigm is made of three layers: input, reservoir and output (see general ESN scheme in **Figure 2A**). In the case of the ESN, the reservoir is a

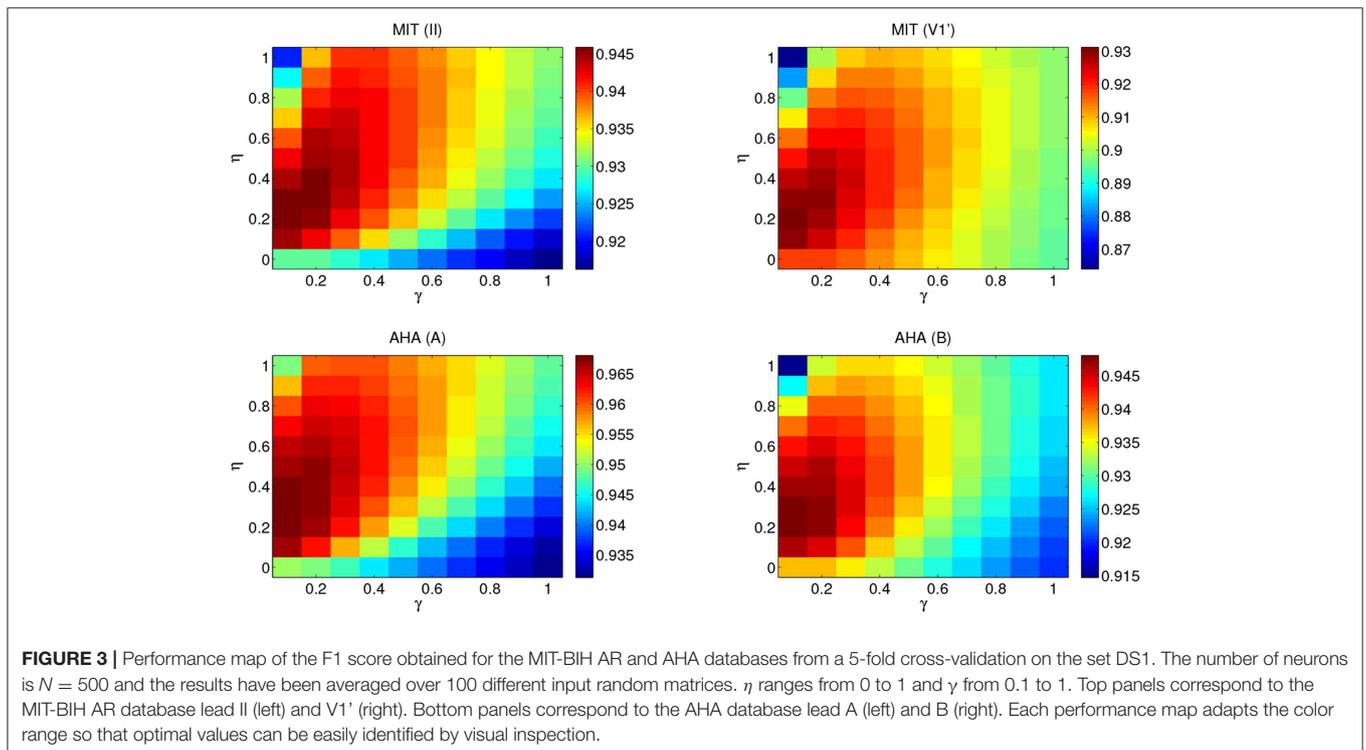
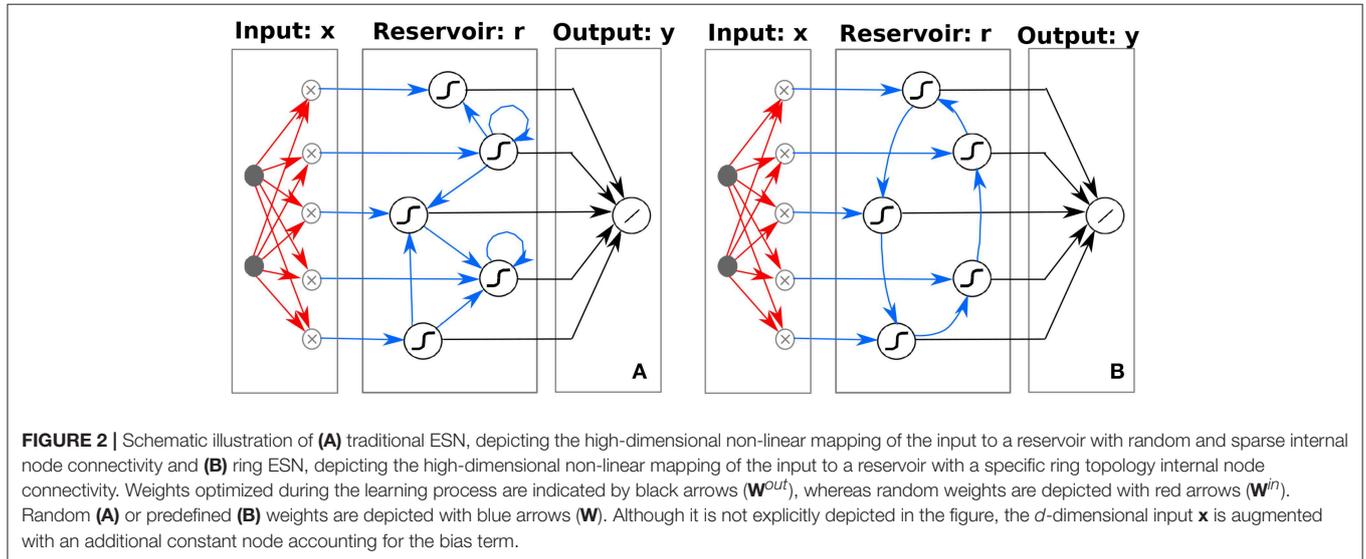
recurrent neural network with random input and random connection weights between the neurons. Thanks to the recurrence of the network, current reservoir responses depend on the previous state of the reservoir, yielding an ESN capable of performing context-dependent computations. The reservoir benefits from a high-dimensional non-linear mapping of the input, so that the reservoir response is easier to classify than the original input by means of a simple linear regression technique.

At the input stage, the ECG data must be fed into the reservoir network. In this process, dimensions must change from  $d \times Hb$

to  $N \times Hb$ , where  $d$ ,  $Hb$ , and  $N$  are the number of input features, heartbeats, and network neurons, respectively. The mapping from the input into the reservoir is done through a random input matrix  $\mathbf{W}_{N \times d}^{in}$  generated from a uniform distribution  $\in [-1, 1]$ . Hence, the ECG data original features vector  $\mathbf{u}_{d \times Hb}$  is modified according to:

$$\mathbf{X}_{N \times Hb} = (\mathbf{W}_{N \times d}^{in} \times \mathbf{u}_{d \times Hb}). \tag{5}$$

Once the first data is fed into the reservoir, the input proceeds sequentially and further reservoir responses are computed



iteratively. The reservoir matrix response  $\mathbf{r}$  for the  $n$ th heartbeat for the standard ESN is obtained as follows:

$$\mathbf{r}(n) = F(\gamma \mathbf{X}(n) + \eta \mathbf{W} \mathbf{r}(n-1)), \quad (6)$$

where  $\mathbf{W}$  is the random connection square matrix, with dimensions  $N \times N$ ,  $F$  is the ESN activation function and  $\gamma$  and  $\eta$  are the input and connection scaling parameters, respectively. For the standard ESN,  $\mathbf{W}$  is also generated from a uniform distribution  $\in [-1, 1]$  and defines the connection weights between the internal neurons. For the non-linear function, we choose the classical sigmoid function with exponent  $-4$  and a bias of  $0.5$ , i.e.,  $F(x) = \frac{1}{1+e^{-4x}} - 0.5$ . Reservoir computers with these sigmoid functions have shown optimal results solving different tasks [35]. Other activation functions, such as rectifiers can also be used.

In this method, only the connections between the reservoir responses and the output are optimized using, usually, some simple linear regression. The response of the ESN to the input,  $\mathbf{r}(n)$ , is used to calculate the expected output,  $\hat{\mathbf{y}}(n)$ , according to:

$$\hat{\mathbf{y}}(n) = \mathbf{W}^{out} \mathbf{r}(n), \quad (7)$$

where  $\mathbf{W}_{l \times N}^{out}$  are the output weights of the ESN and  $l$  the number of output nodes. The output weights are computed by minimizing the squared error between the train outputs and their corresponding target class values, usually employing a linear regression method [36]. In addition, the normal equation formulation is adopted. For the heartbeat classifier we have found that due to the experimental noise present in the original data, simple linear regression results are similar to ridge regression results. For this reason, we prefer the use of linear regression. In this work we deal with a classification task that requires a binary output, e.g., 0 and 1, for the SVEB+ and VEB+ classes, respectively. Thus, the continuous output given by Equation (7)

is converted into a binary one by means of a decision threshold of  $0.5$ .

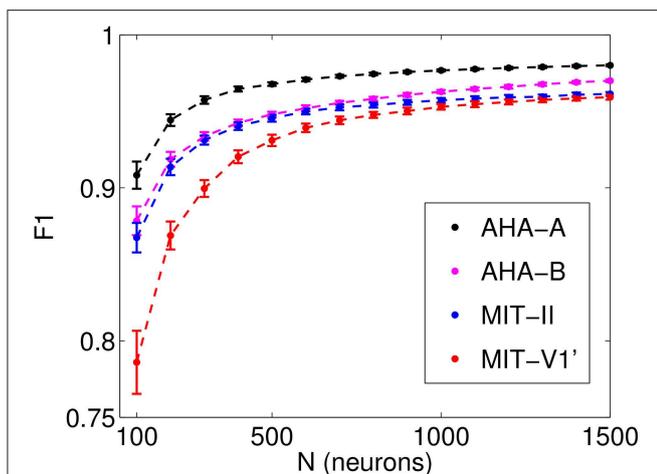
In most of the ESN approaches, the connection matrix  $\mathbf{W}$  is a sparse random matrix. This general form is schematically represented in **Figure 2A**. However, it has recently been shown that simpler ESN with ring topologies perform as well as those with a standard random connection matrix [37]. The ring ESN presents fixed random connections at the input layer  $\mathbf{W}^{in}$  and fixed deterministic weights between internal reservoir neurons, with a connection matrix  $\mathbf{W}$  of only non-zero elements in the lower sub-diagonal  $\mathbf{W}_{i+1,i} = 1$  and at the upper-right corner  $\mathbf{W}_{1,N} = 1$ . The ring ESN is schematically illustrated in **Figure 2B**.

In this work, we use a ESN with ring topology for convenience. The simplicity of the ring ESN allows for an easy exploration of the system parameters in contrast to the computationally demanding trial and error process in ESNs with random topologies [37]. Moreover, this simplicity also allows an easy hardware implementation of the ring ESN using delay-coupled systems [38–41].

## 2.6. Parameter Optimization of the ESN for the SVEB+ and VEB+ Classification

The ring ESN topology allows for a simple optimization procedure, in contrast to the complex trial and error ESN construction with random topologies. The typical model construction decisions in a ring ESN include: setting the network size ( $N$ ), the scaling parameters  $\gamma$  and  $\eta$  and the random input connections ( $\mathbf{W}^{in}$ ). In this heartbeat arrhythmia classification task, the data are very imbalanced [the number of VEB+ cases is much smaller than the SVEB+ ones (see **Table 3**)], and the system is prone to have a high accuracy but a poor classification performance. Thus, the criterion to choose the optimum ring ESN parameters to discriminate between the SVEB+ and VEB+ classes is the one that maximizes the F1 score over the training set DS1.

The optimal  $\eta$  and  $\gamma$  values for each lead and database are determined via a 5-fold cross-validation over the corresponding training set. **Figure 3** shows the performance of the combinations of the pair  $(\eta, \gamma)$  with a fixed number of neurons  $N = 500$  for the MIT-BIH AR and the AHA databases. To avoid an undesired dependence on the sparsity and randomness of the input connections, we average over 100 different input random matrices ( $\mathbf{W}^{in}$ ). The parameter pair that yields the best overall classification is  $\eta = 0.2$  and  $\gamma = 0.1$ . It is worth mentioning that the memory of past heartbeats helps the classification of heartbeats because the case of  $\eta = 0$  (where ESN has no recurrent connections and it is just a feed-forward neural network with one hidden layer) is out of the optimum performance area. This suggests that the memory of past heartbeats helps the classification of present heartbeats. Once the pair  $(\eta, \gamma)$  is set, their optimal values are used to explore the dependence on the number of neurons ( $N$ ) via a 5-fold cross-validation over the corresponding training set. The F1 score as a function of the number of neurons for the value pair  $(\eta = 0.2, \gamma = 0.1)$  is represented in **Figure 4**. As expected, the performance improves with the number of neurons but it starts to saturate for network



**FIGURE 4** | Performance (F1 score) obtained from a 5-fold cross-validation on the set DS1 as a function of the number of neurons ( $N$ ). Results for  $\eta = 0.2$ ,  $\gamma = 0.1$ , which have been averaged over 100 different input random matrices.

sizes over 700 neurons. One of the advantages of the ESN is that they are not prone to overfitting. Performance vs.  $N$  in the test set follows a similar trend than in the training set. We choose a value of  $N = 1,000$  that suits a compromise between good performance for all the studied databases and leads and the computational time. The performance for  $N = 1,000$  is only slightly lower than the one obtained for a larger number of neurons but requires a moderate computational time. The outcome of the optimization must be a fast algorithm suitable for real-time monitoring that, in addition, can be easily retrained when new data are available.

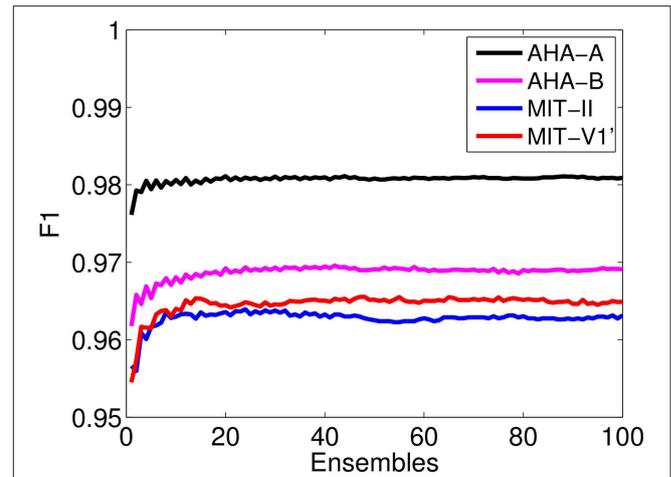
Subsequently, we search for the optimum input connectivity matrix  $\mathbf{W}^{in}$ . A usual approach would be to randomly generate several input matrices and choose the one that performs better in the training set. However, we note that optimizing the input matrix for the training set does not necessarily yield the optimum performance in the test set. Instead, we use a parallel ensemble method in our case since it yields an improvement in the performance. Ensemble methods have already been successfully used for arrhythmia classification [12, 15, 42]. Parallel ensemble methods are learning models that combine the outputs of multiple base classifiers generated in parallel. They exploit the independence between the base classifiers to obtain more accurate predictions than the average error of the individual classifiers. Ensembles are an effective technique if the base classifiers are reasonably accurate and there is diversity between their responses. In an ESN, the mapping of the input data to a high-dimensional non-linear reservoir varies depending on the randomly generated input matrix and this yields variability in the ESN outputs. The output of the ensemble is just the majority voting over the individual outputs of the ESNs. In **Figure 5**, we show the F1 score over the training set DS1 for an ensemble of ESNs with different input matrices as the number of members of the ensembles increases. After combining the outputs of 30 ring ESNs, the classifier performance does not improve when adding new members to the ensemble. Therefore, in the evaluation phase, we use ensembles of 30 ESNs.

In addition, we assess whether a faster alternative to the  $(\eta, \gamma)$  parameter optimization is feasible. To that end, we carry out an ensemble test on a classification that uses random values for the  $(\eta, \gamma)$  reservoir parameters. In this case, each member of the ensemble takes random values for the  $(\eta, \gamma)$  drawn from a uniform distribution between  $[0, 0.8]$  and  $[0.01, 0.5]$  for the  $\eta$  and  $\gamma$  parameters, respectively. Thus, the optimization of  $(\eta, \gamma)$  on the training set would not be necessary. However, we have found that the choice of random  $(\eta, \gamma)$  parameter values is valid for the classification of leads II (MIT-BIH AR) and A (AHA) but it yields a significant decrease in the PPV of leads V1' (MIT-BIH AR) and B (AHA). Therefore,  $\eta = 0.2$  and  $\gamma = 0.1$  are the optimum values used in the Results section.

### 3. RESULTS

#### 3.1. Classifier Evaluation

After optimizing the parameters of the classifier over the training set (DS1) as described in the Methods section, we evaluate the classifier using the optimal parameters. The final



**FIGURE 5** | Performance of the F1 score obtained on the set DS1 as a function of the number of ensembles of ESNs with different input random matrices. The parameters of the ring ESN that form the ensemble are set at their optimal values,  $\eta = 0.2$ ,  $\gamma = 0.1$ , and  $N = 1,000$ .

**TABLE 4** | VEB+ performance over the test set DS2 using an ensemble of 30 ring ESNs.

Database	Lead	Se (%)	PPV (%)	Sp (%)	Acc (%)
MIT-BIH AR	II	84.4 (82.9)	95.8 (85.5)	99.7(98.8)	98.6 (97.7)
	V1'	81.5 (78.9)	76.2 (66.0)	98.0 (96.6)	96.8 (95.3)
AHA	A	90.4 (87.2)	94.9 (92.4)	99.5 (99.2)	98.6 (98.5)
	B	87.9 (85.8)	89.6 (83.4)	98.9 (98.2)	97.8 (97.0)

The values into parenthesis show average of the individual performances of each ring ESN that is part of the ensemble.

performance is evaluated in the test phase with heartbeats that have not been used in the training set and come from different subjects (DS2 set).

**Table 4** shows the classification performance obtained by an ensemble of 30 ring ESNs over the test set DS2. The parameters of the individual ESN are the ones optimized in the training phase. We highlight the fact that the optimal regime for the ESN coincides regardless of database and lead. Since the original heartbeat waveform is normalized between  $[-1, 1]$  and the RR intervals are similar between both databases, the optimum ESN parameters ( $\eta = 0.2$ ,  $\gamma = 0.1$ , and  $N = 1000$ ) coincide for the MIT-BIH AR and the AHA databases. Thus, we expect that these optimum parameters can also be valid for other databases.

The best performance is obtained for the lead A of the AHA database. In the MIT-BIH AR, the lead II gives the best results. Comparing the ensemble results with those obtained with the average of ensemble base classifiers, it is clear that the ensembles reduce the overall error given by a single ESN. The ensembles remarkably reduce the incidence of the false negatives, leading to higher PPV. An ensemble of classifiers has already been used to classify heartbeats and significant improvements have been reported [12, 15]. The improvement in the classification accuracy thanks to the ensembles comes at the cost of higher computation

**TABLE 5** | Cross database VEB+ performance over the test set DS2 using an ensemble of 30 ring ESNs.

Train (DS1)	Test (DS2)	Se (%)	PPV (%)	Sp (%)	Acc (%)
AHA A	AHA A	90.4	94.9	99.5	98.6
	AHA B	87.2	92.4	99.2	98.1
	MIT-BIH AR II	78.2	98.5	99.9	98.3
	MIT-BIH AR V1'	71.5	80.6	98.7	96.7
AHA B	AHA A	82.2	97.1	99.7	98.1
	AHA B	87.9	89.6	98.9	97.8
	MIT-BIH AR II	84.9	97.2	99.8	98.7
	MIT-BIH AR V1'	79.1	43.4	91.9	91.0
MIT-BIH AR II	AHA A	69.4	20.5	71.4	71.2
	AHA B	58.8	23.9	80.0	78.0
	MIT-BIH AR II	84.4	95.8	99.7	98.6
	MIT-BIH AR V1'	39.9	17.5	85.2	81.9
MIT-BIH AR V1'	AHA A	77.0	49.6	91.7	90.3
	AHA B	74.7	49.1	91.8	90.1
	MIT-BIH AR II	72.6	97.6	99.9	97.9
	MIT-BIH AR V1'	81.5	76.2	98.0	96.8

times. However, ensembles are inherently parallel, which can make them much more efficient at training and test time if one has access to a computer with multiple processors.

As part of our study, we assess the generalization capability of our SVEB+ and VEB+ classifier by evaluating the performance of the classifier on a lead and/or database different from the one used to train it. The results are shown in **Table 5**. The best generalization capability is obtained when the classifier is trained either with the AHA lead A or lead B, performing relatively well for all the analyzed leads in the test. The bigger size and the richer variety of the AHA database is likely the reason of the better generalization capability of the classifiers trained with the AHA leads than those trained with the MIT leads. The classification into SVEB+ and VEB+ is based mainly on the morphological shape of the lead. In spite of this lead dependency, the classifier can to some extent generalize to other leads. It is worth mentioning that the MIT-BIH AR cross database performance is relatively poor, specially for the lead II. Some ECG recordings of MIT-BIH AR lead V1' are V2 or V5, which could lead to a better generalization capability of the lead V1' but also to a worse performance in the intra-lead classification when compared with the other intra-lead performances (see **Table 4**).

### 3.2. Computational Times

Besides providing a detailed characterization of the arrhythmia heartbeat classifier based on ESNs, our study also aims at achieving computational times that allow for real-time processing of ECG data. In particular, we have implemented the ESN classifiers described here independently in an unparallelized C++ version for the CPU and a C++/CUDA version for the GPU. C++ refers to the object oriented programming language and CUDA is a parallel computing platform developed by the company Nvidia to interface with their GPUs. The

**TABLE 6** | Technical specifications of the CPU and GPU used in this work.

	CPU	GPU
Processor	Intel(R) Core(TM) i7-4790K	NVIDIA TITAN X Pascal (3584 CUDA cores)
Clock frequency	4,400 MHz	1,417 MHz
Memory	32 GB	12 GB
Max. Mem. Bandwidth	25.6 GB/s	480 GB/s

specific technical details for the CPU and GPU are summarized in **Table 6**.

Although ensembles are inherently independent, making them good candidates for parallel multi-processor implementations, the presence of large matrix products and non-linear mapping functions in the reservoir paradigm also makes serial implementations suitable for the exploration of computationally fast approaches. These approaches, such as GPU implementations, are capable of reducing the latency and increasing the throughput.

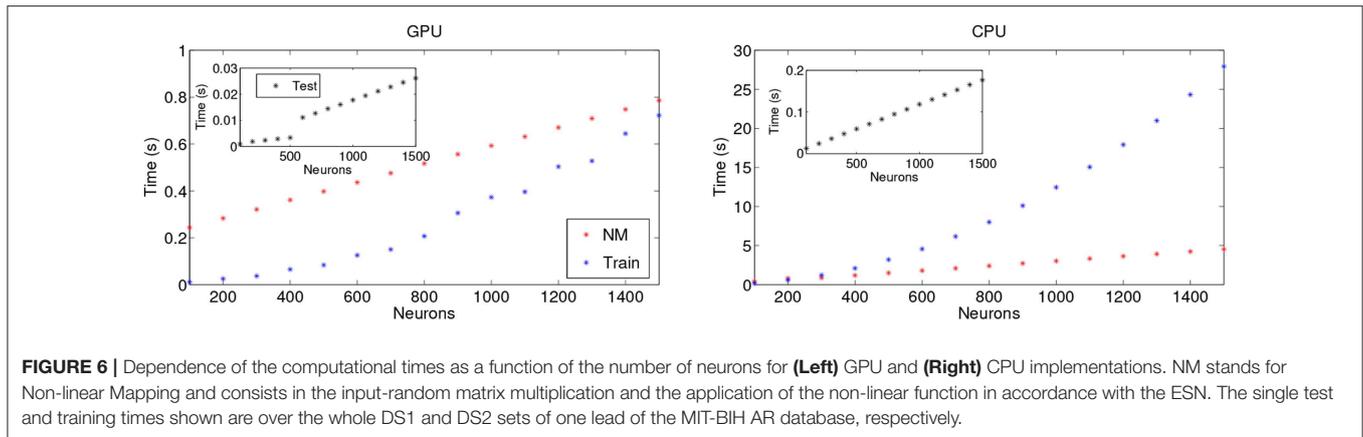
In order to explore the computational time and reservoir size (N) dependence, a series of training and classification procedures for the MIT-BIH AR database are analyzed. Linear regressions are carried out by means of lower-upper decomposition. C++ implementations benefit from the Eigen library<sup>1</sup>, while C++/CUDA use cuSolver, cuBLAS products and a CUDA kernel implemented for the non-linear mapping.

**Figure 6** shows the computational times of a training and a testing realization for the DS1 and DS2 sets of the MIT-BIH AR databases, respectively, vs. the number of neurons. The GPU and CPU comparison highlights the advantage of using a GPU implementation, with significantly lower training times. The depicted computational times include, on the one hand, the random non-linear mapping of the input onto the reservoir and, on the other hand, the calculation of the output weights  $W^{out}$  over the entire train dataset. The insets in **Figure 6** show the computational time for the final classification product steps that calculate the output in the test dataset. As expected, the processing time increases with the number of neurons, especially in the training procedure. The influence of small sized products on cuBLAS scaling, intrinsic to the library, can be seen in the piece-wise linear trend present in the GPU Classification product. The reported computational times account for 11 h of ECG recordings, allowing the exploration of different parameter regimes and providing fast classifications clearly suitable for real-time scenarios that may include statistical ensembles.

### 3.3. Comparison With Other Heartbeat Arrhythmia Classifiers

The MIT-BIH AR database is by far the most used to evaluate methods on the ventricular arrhythmia classification. However, making a fair comparison between heartbeat classifiers is a difficult task. For instance, classifiers sharing heartbeats for the

<sup>1</sup>Eigen v3.3—Gaël Guennebaud, Benoît Jacob et al.  
<http://eigen.tuxfamily.org/>



**FIGURE 6 |** Dependence of the computational times as a function of the number of neurons for **(Left)** GPU and **(Right)** CPU implementations. NM stands for Non-linear Mapping and consists in the input-random matrix multiplication and the application of the non-linear function in accordance with the ESN. The single test and training times shown are over the whole DS1 and DS2 sets of one lead of the MIT-BIH AR database, respectively.

**TABLE 7 |** VEB performance of the heartbeat classifiers on the MIT-BIH AR database.

Work	Feature set	Classifier	Leads	VEB	
				Se	PPV
de Chazal et al. [3]	Morphological, RR-intervals	Weighted LD	II + V1'	77.7	81.9
Ye et al. [18]	Morphological, RR-intervals, wavelet, ICA, PCA	SVM	II + V1'	81.5	63.1
Zhang et al. [14]	Morphological, RR-intervals	Feature selection + SVM	II + V1'	85.5	92.8
Mar et al. [4]	Morphological, HOS, temporal features	Feature selection + MLP	II + V1'	86.8	75.9
Garcia et al. [16]	Morphological, wavelets, TVCG	PSO + SVM	II + V1'	87.3	59.4
Llamedo and Martinez [31]	Morphological, RR-interval, VCG, wavelet	LD+ EMC	II + V1'	83.0	88.0
Llamedo and Martinez [33]	RR-interval, wavelet	LD+ EMC	II + V1'	89.0	87.0
Ye et al. [26]	Morphological, RR-intervals, wavelet, ICA	General + specific classification model	II + V1'	91.8	98.0
Tejairo et al. [43]	Morphological, rhythm features, RR-intervals	Abductive interpretation	II + V1'	94.6	96.8
Ghorbani et al. [7]	Morphological, RR-intervals, statistical features, GMM + EM	Decision trees	II + V1'	96	77.6
Krasteva et al. [5]*	Morphological, RR-intervals, correlations	Decision trees	II + V1'	96.7	99.2
Wu et al. [20]	DBN, RR-intervals	Softmax regression	II	80.5	81.4
Lannoy et al. [19]	Morphological, RR-intervals, HOS, HBF coeff	Weighted conditional random fields	II	85.1	–
Rahhal et al. [22] *	Raw ECG data	Deep neural networks	II	91.0	79.5
Raj et al. [17]	DOST	PSO + SVM	II	87.5	65.4
Sultan Qurraie and Ghorbani Afkhami [6]	RR-interval, HOS, time–frequency	Decision trees	II	95.4	94.1
Herry et al. [44]	RR-interval, SST	SVM	II V1'	77.5 79.6	79.1 62.7
Huang et al. [15]	Random projections	SVM ensembles	II V1'	93.9 78.1	90.9 43.8
This work	Raw ECG data, RR-intervals	ESN ensembles	II V1'	92.7 86.1	95.7 75.1

Only the best fully automatic work result is reported. All the classifiers have been trained over the set DS1 and tested over DS2, except the ones marked with\*. Rahhal et al. [22] and Krasteva et al. [5] test against all the MIT-BIH AR database. Rahhal et al. [22] trains over the DS1 and Krasteva et al. [5] uses three databases (AHA, MIT-BIH-SV, and EDB) to train the model. See the text for a description of the different methods and features.

same subjects in the training and test set have unrealistically better evaluation results than classifiers that follow the inter-patient procedure [7]. Semi-automatic heartbeat classifiers (that require some assistance for expert cardiologist) also have a better performance than the fully automatic approaches [33]. Thus, to be as fair as possible, we only compared our method with other fully automatic heartbeat classifiers that make the test over the DS2 set of the MIT-BIH AR database and whose train set does not share subjects with the testing set.

Focusing on the detection of ventricular arrhythmia, we compare the VEB (V) performance instead of the VEB+ (V+F), as the VEB+ performance is usually not reported in the literature. The VEB performance has then been calculated in our algorithm without taking into account the F heartbeats, which are rather rare. **Table 7** compares the VEB detection performance of state-of-the-art algorithms with the method proposed in this manuscript. **Table 7** also provides information about the features and classifiers used by the different approaches. In most cases,

the computational cost of these methods, either during the training or the test phases, is not mentioned. **Table 7** presents a wide variety of methods, such as Multilayer Perceptron (MLP), temporal vectorcardiogram (TVCG), Expectation-maximization clustering algorithm (EMC), Gaussian mixture modeling (GMM), Enhanced expectation maximization (EM), Orthogonal Stockwell Transform (DOST), Deep Belief Networks (DBN), and synchrosqueezing transform (SST).

Our method outperforms or shows state-of-the-art results with methods that used much more complicated procedures to extract and select the heartbeat features for the VEB class. Some of the methods with better performance than the method proposed here are not well-suited for real-time applications, as the feature extraction stage can not be implemented online, such as in [43] or imply a high computational cost [6]. Moreover, our approach outperforms the other single lead classifiers reported for the VEB classification based on the MIT-BIH AR lead V1', showing a better generalization capability than the other methods based on a single lead. Finally, the excessive false alarm rate (low PPV) is a major problem for clinical use since it diminishes the confidence in the algorithm. The approach discussed in this manuscript has the best PPV for the VEB class among the single lead classifiers.

## 4. DISCUSSION

The proposed method shows excellent classification results for the VEB class on the MIT-BIH AR and the AHA databases, outperforming existing single lead classification algorithms in the detection of ventricular arrhythmia. The presented ESN approach is suitable for processing long-term recordings and large databases as the feature extraction and the algorithm itself both have minimal computational requirements.

Overall, the ESN presents two main advantages over other classical methods that have been used to classify heartbeats, such as the SVM, NN, and decision trees (see **Table 7**). First, the aforementioned methods involve relatively time consuming complex computations in the training phase that in ESN are easily computed. We have checked that the computation times of the classification algorithm for the evaluation of 11 h of ECG recordings amounts to <0.2 s for a lab CPU, while the use of a GPU (see **Table 6**) offers at least a speedup of an order of magnitude. Second, past heartbeats play a role in the classification task in the case of the ESN thanks to its intrinsic memory, having a positive impact on the performance.

## REFERENCES

- Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PC, Mark RG, et al. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation*. (2000) **101**:e215–20. doi: 10.1161/01.cir.101.23.e215
- da S Luz EJ, Schwartz WR, Cámara-Chávez G, Menotti D. ECG-based heartbeat classification for arrhythmia detection: a survey. *Comput Methods Programs Biomed*. (2016) **127**:144–64. doi: 10.1016/j.cmpb.2015.12.008

In this work, heartbeats are classified as SVEB+ and VEB+. Future work will focus on the extension of these results to the five heartbeat classes recommended by the AAMI. Another important aspect not covered in our study is the fixed heartbeat window length that can be inappropriate in the case of fast and slowly varying heart rhythms when changing physical activity. Thus, there is a need to study adaptive beat size segmentation. The understanding of the exact relation between underlying physiology and features is a potential question to address. However, there are no conclusive guidelines about which features should be used to diagnose arrhythmias from the ECG using computer aided systems.

## DATA AVAILABILITY

The MIT-BIH AR publicly available dataset was part of the analysis presented in this work. This database can be found here: <https://physionet.org/physiobank/database/mitdb/>.

## AUTHOR CONTRIBUTIONS

MA implemented the classifier and performed the computational realizations. SO and MS designed and supervised the project. All authors contributed to the discussion of the results and to the writing of the manuscript.

## FUNDING

Authors gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan X Pascal GPU used for this research. This work was partially funded by the Spanish Ministerio de Economía y Competitividad (MINECO) and Fondo Europeo de Desarrollo Regional (FEDER) and the European Social Fund through project TEC2016-80063-C3-3-R (MINECO/AEI/FEDER/UE). MA was supported by the Beca de colaboración 012/2016 UIB fellowship on Information processing in neural and photonic systems. MS was supported by the Spanish Ministerio de Economía, Industria y Competitividad through a Ramón y Cajal Fellowship (RYC-2015-18140). SO was supported by the Conselleria d'Innovació, Recerca i Turisme del Govern de les Illes Balears and the European Social Fund.

## ACKNOWLEDGMENTS

The authors would like to thank Claudio Mirasso, Ingo Fischer, Xavier Ibáñez Català, and Agustín Macià for valuable scientific discussions.

- de Chazal P, O'Dwyer M, Reilly RB. Automatic classification of heartbeats using ECG morphology and heartbeat interval features. *IEEE Trans Biomed Eng*. (2004) **51**:196–206. doi: 10.1109/TBME.2004.827359
- Mar T, Zaunseder S, Martínez JP, Llamedo M, Poll R. Optimization of ECG classification by means of feature selection. *IEEE Trans Biomed Eng*. (2011) **58**:2168–77. doi: 10.1109/TBME.2011.2113395
- Krasteva V, Jekova I, Leber R, Schmid R, Abacherli R. Superiority of classification tree versus cluster, fuzzy and discriminant models

- in a heartbeat classification system. *PLoS ONE*. (2015) **13**:e0140123. doi: 10.1371/journal.pone.0140123
6. Sultan Qurraie S, Ghorbani Afkhami R. ECG arrhythmia classification using time frequency distribution techniques. *Biomed Eng Lett*. (2017) **7**:325–32. doi: 10.1007/s13534-017-0043-2
  7. Ghorbani Afkhami R, Azarnia G, Tinati MA. Cardiac arrhythmia classification using statistical and mixture modeling features of ECG signals. *Pattern Recognit Lett*. (2016) **70**:45–51. doi: 10.1016/j.patrec.2015.11.018
  8. Dokur Z, Ölmez T. ECG beat classification by a novel hybrid neural network. *Comput Methods Programs Biomed*. (2001) **66**:167–81. doi: 10.1016/S0169-2607(00)00133-4
  9. Elhaj FA, Salim N, Harris AR, Swee TT, Ahmed T. Arrhythmia recognition and classification using combined linear and nonlinear features of ECG signals. *Comput Methods Programs Biomed*. (2016) **127**:52–63. doi: 10.1016/j.cmpb.2015.12.024
  10. Martis RJ, Acharya UR, Min LC. ECG beat classification using PCA, LDA, ICA and discrete wavelet transform. *Biomed Signal Process Control*. (2013) **8**:437–48. doi: 10.1016/j.bspc.2013.01.005
  11. Inan OT, Giovangrandi L, Kovacs GTA. Robust neural-network-based classification of premature ventricular contractions using wavelet transform and timing interval features. *IEEE Trans Biomed Eng*. (2006) **53**:2507–15. doi: 10.1109/TBME.2006.880879
  12. Javadi M, Ebrahimpour R, Sajedin A, Faridi S, Zakernejad S. Improving ECG classification accuracy using an ensemble of neural network modules. *PLoS ONE*. (2011) **6**:e24386. doi: 10.1371/journal.pone.0024386
  13. Kiranyaz S, Ince T, Gabbouj M. Real-time patient-specific ECG classification by 1-D convolutional neural networks. *IEEE Trans Biomed Eng*. (2016) **63**:664–75. doi: 10.1109/TBME.2015.2468589
  14. Zhang Z, Dong J, Luo X, Choi KS, Wu X. Heartbeat classification using disease-specific feature selection. *Comput Biol Med*. (2014) **46**:79–89. doi: 10.1016/j.cmpbiomed.2013.11.019
  15. Huang H, Liu J, Zhu Q, Wang R, Hu G. A new hierarchical method for inter-patient heartbeat classification using random projections and RR intervals. *Biomed Eng Online*. (2014) **13**:90. doi: 10.1186/1475-925X-13-90
  16. Garcia G, Moreira G, Menotti D, Luz E. Inter-patient ECG heartbeat classification with temporal VCG optimized by PSO. *Sci Rep*. (2017) **7**:10543. doi: 10.1038/s41598-017-09837-3
  17. Raj S, Ray KC, Shankar O. Cardiac arrhythmia beat classification using DOST and PSO tuned SVM. *Comput Methods Programs Biomed*. (2016) **136**:163–77. doi: 10.1016/j.cmpb.2016.08.016
  18. Ye C, Kumar BVK, Coimbra MT. Heartbeat classification using morphological and dynamic features of ECG signals. *IEEE Trans Biomed Eng*. (2012) **59**:2930–41. doi: 10.1109/TBME.2012.2213253
  19. De Lannoy G, François D, Delbeke J, Verleysen M. Weighted conditional random fields for supervised interpatient heartbeat classification. *IEEE Trans Biomed Eng*. (2012) **59**:241–7. doi: 10.1109/TBME.2011.2171037
  20. Wu Z, Ding X, Zhang G. A novel method for classification of ECG arrhythmias using deep belief networks. *Int J Comput Intell Appl*. (2016) **15**:1650021. doi: 10.1142/S1469026816500218
  21. Acharya UR, Oh SL, Hagiwara Y, Tan JH, Adam M, Gertych A, et al. A deep convolutional neural network model to classify heartbeats. *Comput Biol Med*. (2017) **89**:389–96. doi: 10.1016/j.cmpbiomed.2017.08.022
  22. Rahhal MMA, Bazi Y, Alhichri H, Alajlan N, Melgani F, Yager RR. Deep learning approach for active classification of electrocardiogram signals. *Inf Sci*. (2016) **345**:340–54. doi: 10.1016/j.ins.2016.01.082
  23. Li H, Yuan D, Ma X, Cui D, Cao L. Genetic algorithm for the optimization of features and neural networks in ECG signals classification. *Sci Rep*. (2017) **7**:41011. doi: 10.1038/srep41011
  24. Ortín S, Soriano MC, Alfaras M, Mirasso CR. Automated real-time method for ventricular heartbeat classification. *Comput Methods Programs Biomed*. (2019) **169**:1–8. doi: 10.1016/j.cmpb.2018.11.005
  25. Zidelmal Z, Amirou A, Ould-Abdeslam D, Merckle J. ECG beat classification using a cost sensitive classifier. *Comput Methods Programs Biomed*. (2013) **111**:570–7. doi: 10.1016/j.cmpb.2013.05.011
  26. Ye C, Kumar BVK, Coimbra MT. An automatic subject-adaptable heartbeat classifier based on multiview learning. *IEEE J Biomed Health Inform*. (2016) **20**:1485–92. doi: 10.1109/JBHI.2015.2468224
  27. Jaeger H, Haas H. Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication. *Science*. (2004) **304**:78–80. doi: 10.1126/science.1091277
  28. Moody GB, Mark RG. The impact of the MIT-BIH arrhythmia database. *IEEE Eng Med Biol Mag*. (2001) **20**:45–50. doi: 10.1109/51.932724
  29. Emergency Care Research Institute. *American Heart Association (AHA) Ventricular Arrhythmia ECG Database*. Plymouth Meeting, PA: Emergency Care Research Institute (2003).
  30. ANSI/AAMI. *Testing and Reporting Performance Results of Cardiac Rhythm and ST Segment Measurement Algorithms*. Arlington, VA: American National Standards Institute, Inc (ANSI), Association for the Advancement of Medical Instrumentation (AAMI), ANSI/AAMI/ISO EC57, 1998-(R)2008 (2008).
  31. Llamedo M, Martinez JP. Heartbeat classification using feature selection driven by database generalization criteria. *IEEE Trans Biomed Eng*. (2011) **58**:616–25. doi: 10.1109/TBME.2010.2068048
  32. Martinez JP, Almeida R, Olmos S, Rocha AP, Laguna P. A wavelet-based ECG delineator: evaluation on standard databases. *IEEE Trans Biomed Eng*. (2004) **51**:570–81. doi: 10.1109/TBME.2003.821031
  33. Llamedo M, Martinez JP. An automatic patient-adapted ECG heartbeat classifier allowing expert assistance. *IEEE Trans Biomed Eng*. (2012) **59**:2312–20. doi: 10.1109/TBME.2012.2202662
  34. Lukoševičius M, Jaeger H. Reservoir computing approaches to recurrent neural network training. *Comput Sci Rev*. (2009) **3**:127–49. doi: 10.1016/j.cosrev.2009.03.005
  35. Ortín S, Pesquera L. Reservoir computing with an ensemble of time-delay reservoirs. *Cognit Comput*. (2017) **9**:327–36. doi: 10.1007/s12559-017-9463-7
  36. Lukoševičius M. A practical guide to applying echo state networks. In: Montavon G, Orr GB, Müller KR, editors. *Neural Networks: Tricks of the Trade*. 2nd ed. Berlin; Heidelberg: Springer Berlin Heidelberg (2012). p. 659–86. doi: 10.1007/978-3-642-35289-8\_36
  37. Rodan A, Tino P. Minimum complexity echo state network. *IEEE Trans Neural Netw*. (2011) **22**:131–44. doi: 10.1109/TNN.2010.2089641
  38. Appeltant L, Soriano MC, Van der Sande G, Danckaert J, Dambre J, Schrauwen B, et al. Information processing using a single dynamical node as complex system. *Nat Commun*. (2011) **2**:468. doi: 10.1038/ncomms1476
  39. Paquot Y, Dupont F, Smerieri A, Dambre J, Schrauwen B, Haelterman M, et al. Optoelectronic reservoir computing. *Sci Rep*. (2012) **2**:287. doi: 10.1038/srep00287
  40. Brunner D, Soriano MC, Mirasso C, Fischer I. Parallel photonic information processing at gigabyte per second data rates using transient states. *Nat Commun*. (2013) **4**:1364. doi: 10.1038/ncomms2368
  41. Ortín S, Soriano MC, Pesquera L, Brunner D, San-Martín D, Fischer I, et al. A unified framework for reservoir computing and extreme learning machines based on a single time-delayed neuron. *Sci Rep*. (2015) **5**:14945. doi: 10.1038/srep14945
  42. Osowski S, Linh TH. ECG beat recognition using fuzzy hybrid neural network. *IEEE Trans Biomed Eng*. (2001) **48**:1265–71. doi: 10.1109/10.959322
  43. Teijeiro T, Felix P, Presedo J, Castro D. Heartbeat classification using abstract features from the abductive interpretation of the ECG. *IEEE Journal of Biomedical and Health Inform*. (2018) **22**:409–20. doi: 10.1109/JBHI.2016.2631247
  44. Herry CL, Frasch M, Seely AJE, Wu HT. Heart beat classification from single-lead ECG using the synchrosqueezing transform. *Physiol Meas*. (2017) **38**:171–87. doi: 10.1088/1361-6579/aa5707

**Conflict of Interest Statement:** MA is currently employed by company PLUX S.A. This research was entirely conducted while he was an IFISC researcher.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Alfaras, Soriano and Ortín. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.