Detection of Atrial Fibrillation From ECG Using BTD Tensor Decomposition

Renan H. Cardoso, C. Alexandre R. Fernandes, Pedro M. R. de Oliveira

***Abstract*—Atrial fibrillation (AF) is a common cardiac arrhythmia associated with various cardiovascular diseases and has a significant impact on mortality around the world. This work focuses on the detection of AF using data collected from cardiac monitoring through the Electrocardiogram (ECG), proposing new attributes for the prediction of AF. In particular, the present work proposes novel convergence and optimization indicators derived from Block-term Decomposition (BTD), applied to five RRI ECG segments, combined with RRI intervals (RRI) to improve the detection of AF. These features were used with tree-based machine learning algorithms to classify signals as Atrial Fibrillation (AF) or Normal Sinus Rhythm (NSR). The study also discusses data acquisition from three different ECG databases: Atrial Fibrillation Database (AFDB), Long-term Atrial Fibrillation Database (LTAFDB), and Normal Sinus Rhythm Database (NSRDB).**

***Index Terms*—Atrial Fibrillation, Tensor Decomposition, Electrocardiogram, Automatic Detection, Machine Learning**

I. Introduction

BASED on data from the Global Burden of Disease (GBD) 2019 study [36], the global prevalence of Atrial Fibrillation (AF) has escalated to 59.7 million cases, reflecting a dramatic rise since 28.3 million in 1990 to 45.6 million in 2010 [37]. Although the prevalence of undiagnosed AF in the community remains imprecise, backcalculation suggest that nearly 11% (around 591.000 cases) out of more than 5.6 million AF occurrences in the United States remained undiagnosed as of 2015 [38].

The total burden measured in disability adjusted life years (DALYs) attributable to AF has similarly surged, increased from 3.8 million in 1990 to 8.4 million in 2019 [39]. Moreover, the impact of modifiable risk factors, including hypertension, alcohol consumption, obesity, and smoking, on DALY related to AF has notably escalated during this period [36].

Although the electrophysiological underpinnings of AF remain partially elusive, extensive research in recent years has continued to shed light on various aspects of the condition. Typically, AF is suspected based on symptoms including dyspnea, chest discomfort, and irregular pulse patterns [7]. In addition, cardiologists examine Electrocardiogram (ECG) signals detect subtle features indicative of AF.

Employing computational methods, particularly machine learning algorithms [8], for analyzing ECG signals in the context of AF detection has shown significant potential. This advancement has markedly improved both the precision and rapidity of diagnostic outcomes, underpinning prompt intervention and treatment optimization.

Alternatively, techniques based on tensor decompositions applied to ECG data have been widely used over the past decade as a fundamental tool in Blind Source Separation (BSS), particularly in the analysis of atrial activity [9], [11], [21]. Notably, tensor decomposition methods, such as Block-term Decomposition (BTD), have emerged as robust alternatives to traditional matrix approaches by offering unique advantages. These methods can circumvent the limitations of matrix techniques, ensuring uniqueness under broader conditions [11]. Moreover, the temporal robustness of BTD underlines the benefits of employing tensor-based strategies in such analyses [21]. In summary, the use of BTD for the evaluation of ECG signals offers a promising strategy for extracting and interpreting Atrial Activity (AA), particularly in the BSS realm.

So far, tensor-based methods applied to ECG signals, particularly using BTD, have been largely confined to source separation tasks, with clinical evaluations focusing on sources identified with AF. Here, we introduce a novel paradigm where the convergence parameters derived from Hankel-BTD are harnessed, not for source separation, but as informative features for enhancing AF detection. Thus, these parameters serve as key input features for machine learning classifiers targeting AF detection.

The rationale for this approach comes from the observation that the Hankel-BTD tensor decomposition applied to ECG recordings yields unique results primarily in the presence of AF. This uniqueness arises because AF signals typically exhibit atrial components that conform to all-pole (exponential) modeling. Consequently, the inherent uniqueness properties of the Hankel-BTD decomposition are maintained.

In contrast, R-R intervals (RRI), representing the periods between consecutive R waves in the ECG signal, serve as a robust metric for detecting rhythm irregularities linked to AF, have been extensively utilized in the literature to address the challenges of AF detection [12]. Combined, the fusion of BTD convergence parameters and RRIs offers compelling discriminatory power for reliable AF detection.

The primary objective of this work is to introduce an innovative methodology for the detection of AF from ECG recordings by synchronizing the Hankel-BTD convergence

Renan H. Cardoso is with the Federal University of Ceara, Campus Sobral, Ceará, Brazil (e-mail: renanhc@alu.ufc.br). ORCID: 0009-0001-1440-1177.

C. Alexandre R. Fernandes is with the Department of Computer Engineering, Federal University of Ceara, Sobral, Brazil, e-mail: alexandrefernandes@ufc.br. ORCID: 0000-0002-9933-9930.

Pedro Marinho R. de Oliveira is with SOS Oxygène, France (e-mail: pedro.marinho@sosoxygene.com). ORCID: 0000-0002-5310-4714.

This work was supported in part by FUNCAP / Brazil under Grant BP5-0197-00183.01.06/23.

Submission: 2024-05-27, First decision: 2024-08-21, Acceptance: 2025-05-20, Publication: 2025-05-28.

Digital Object Identifier: 10.14209/jcis.2025.3

parameters with the RRI, thus improving the classification between AF and Normal Sinus Rhythm (NSR). This marks a pioneering utilization of tensor decomposition tailored specifically for AF classification. In addition, the study aims to develop a segmented patient data repository derived from three prominent ECG databases (AFDB, NSRDB, LTAFDB) and evaluate the performance of tree-based machine learning classifiers to increase the accuracy, sensitivity, and specificity of AF detection in continuous ECG monitoring. Collectively, this approach sets a new benchmark for ECG analysis by leveraging unique BTD-derived features to significantly enhance AF detection performance.

The rest of the work is organized as follows. Section II reviews traditional and modern approaches for AF detection, highlighting the novelty of using BTD not for source separation but as a feature extraction method. In Section III, the mathematical foundation of BTD and the Hankel-BTD model are presented, detailing how Hankel matrices are used to extract atrial activity. In Section IV, the three ECG databases used in the work are described. AFDB, LTAFDB, and NSRDB. Section V outlines the study methodology, including data preparation, segmentation, feature extraction using BTD convergence parameters, and classifiers. In Section VI, the results of the cross-validation and the test set using five tree-based classifiers are presented. Finally, Section VII presents the main conclusions of the work and the perspectives for future work.

II. Related Works

Automated detection of AF in ECG signals has been a focus of research for several decades, evolving from early statistical models to machine learning techniques. Foundational studies such as those of [42] and [12] introduced time-domain approaches including Hidden Markov Models and RRI-based histogram analysis, both of which demonstrated strong classification performance on benchmark datasets such as AFDB. Later, [43] proposed an innovative image-based technique using ΔRR patterns plotted in grid panels, achieving high classification accuracy through LSVM. Complementing these efforts, [44] and [50] expanded the analysis into heart rate variability and frequency domain transforms, revealing that statistical and spectral markers carry a critical diagnostic value to identify AF episodes.

Building upon these classical methods, more recent studies have explored hybrid frameworks to better manage noisy signals and complex AF morphologies. For example, [46] and [47] introduced methodologies that merge time-domain analysis with wavelet transforms, Lorenz plots, and multi-feature extraction, including atrial activity signals. These approaches not only improved robustness, but also demonstrated strong generalization across varied data sources. Similarly, [25] proposed a probabilistic density histogram of RR intervals as a unified statistical feature, which, when combined with an SVM classifier, shows significant performance gains in terms of accuracy, sensitivity, and specificity.

However, deep learning techniques, especially Convolutional Neural Networks (CNNs), have shown promise in the automated detection of AF. Studies, such as [32], combine CNNs with Recurrent Neural Networks (RNNs), achieving sensitivity and specificity rates that exceed 98% in different databases. Currently, these systems have demonstrated superior performance compared to traditional statistical methods, highlighting the potential of deep learning to enhance AF diagnosis [51]. In addition, new techniques are being developed to suppress overfitting [52], a common obstacle in complex models, characterized by excessive adaptation to training data, capturing noise, and data set-specific patterns. These techniques can be applied to compressed ECG data without the need for reconstruction [53]. Finally, there is growing interest among researchers in applications using this new approach, as evidenced in [54]–[56], paving the way for improved clinical outcomes in the treatment of cardiovascular diseases.

Alongside these approaches, classic BSS methods have been widely used to extract AA from ECG signals. In [16], the Independent Component Analysis (ICA) technique was used to identify independent sources of cardiac activity, validating the separation based on signal kurtosis and spectral analysis. In a subsequent advancement, [28] proposed a BSS modeling approach using a low-rank BTD, which demonstrated superiority in signal extraction in noisy scenarios, especially for AA.

Furthermore, the work [29] was the pioneer in non-invasive extraction of AA during AF episodes using Hankel-BTD tensor decomposition. In 2019, [22] discussed the characteristics of persistent AF, highlighting the challenges in extracting AA due to disorganized and low-amplitude signals. To overcome these limitations, the research utilized the Löwner matrix structure, resulting in improvements in the quality of the AA extraction compared to traditional matrix-based approaches, such as Principal Component Analysis (PCA) and RobustPCA-f.

Recently, [30] introduced the Alternating Group Lasso (AGL) algorithm to compute the estimated tensor through BTD, addressing the limitations of the Alternating Least Squares - Enhanced Line Search (ALS-ELS) algorithm with respect to the choice of structural parameters. This new method demonstrated superiority in robustness and efficiency, allowing for a more precise extraction of AA.

Furthermore, [31] proposed a new approach, called LCAGL, which uses the Löwner matrix structure together with the CAGL algorithm. This strategy focuses on extracting the Atrial Voltage (AV) due to the challenges associated with the low amplitude of AA during persistent episodes of AF. The results indicated successful extraction of AA even in short signal segments, demonstrating the potential of this technique for AF analysis.

In summary, existing research highlights the relevance of extracting and analyzing AA for more accurate diagnoses of AF, highlighting the need to integrate robust signal analysis methods and machine learning. This work aims to use this accumulated knowledge to extract attributes from ECG signals and subsequently classify them, with the aim of improving AF detection.

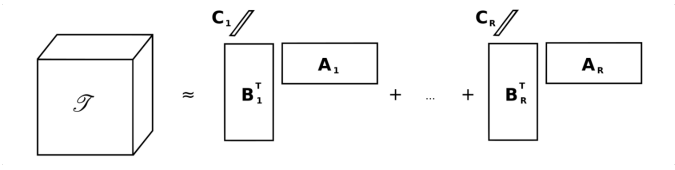


Fig. 1. Visual representation of the BTD with multilinear rank $(L_r, L_r, 1)$.

III. Tensor Prerequisites

The BTD, introduced in [14], represents a tensor $\mathcal{T} \in \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_N}$ as a sum of tensors of lowest order. The BTD can also be viewed as a multilinear rank decomposition $(L_r, L_r, 1)$, which has attracted greater interest due to its more frequent occurrence in applications and the existence of more relaxed uniqueness conditions. Let $\mathcal{T} \in \mathbb{R}^{I \times J \times K}$ be a third order tensor, the BTD with rank $(L_r, L_r, 1)$ is described below:

$$\mathcal{T} = \sum_{r=1}^{R} \mathbf{E}_r \circ \mathbf{c}_r, \tag{1}$$

where $\circ$ is the outer product, $\mathbf{E}_r \in \mathbb{R}^{I \times J}$ is a matrix with rank L_r and $\mathbf{c}_r \in \mathbb{R}^K$ is a vector. Performing a factorization on $\mathbf{E}_r$, results in $\mathbf{A}_r \mathbf{B}_r^T$, where $\mathbf{A}_r \in \mathbb{R}^{I \times L}$ and $\mathbf{B}_r \in \mathbb{R}^{J \times L}$, as follows:

$$\mathcal{T} = \sum_{r=1}^{R} (\mathbf{A}_r \mathbf{B}_r^T) \circ \mathbf{c}_r. \tag{2}$$

The visual representation of the BTD is shown in Fig. 1.

Hankel-BTD, also known as Hankel-based BTD, proposed in [28], is a technique for BSS that uses the BTD of a tensor composed of Hankel matrices. Consider that R source signals with N time samples are organized into a matrix denoted by $\mathbf{S} \in \mathbb{R}^{R \times N}$. In a BSS model, the observed signals are modeled as follows:

$$\mathbf{Y} = \mathbf{MS} \in \mathbb{R}^{K \times N}, \tag{3}$$

where $\mathbf{Y} \in \mathbb{R}^{K \times N}$ is a matrix with the observed signals, $\mathbf{M} \in \mathbb{R}^{K \times R}$ is the mixing matrix and K is the number of observed signals. The objective of the BSS is to estimate $\mathbf{M}$ and $\mathbf{S}$ from the matrix $\mathbf{Y}$.

In Hankel-BTD, a tensor $\mathcal{H} \in \mathbb{R}^{I \times J \times K}$ is built from N Hankel matrices $\mathbf{H}_\mathbf{Y}^{(n)} \in \mathbb{R}^{I \times J}$ constructed from the rows of $\mathbf{Y}$. Each element (i, j) of the hankelized matrix $\mathbf{H}_\mathbf{Y}^{(n)}$ is represented by the following mapping:

$$h_{i,j}^{(n)} = y_{n,i+j-1} \tag{4}$$

where $i = 1, 2, \ldots, I$, $j = 1, 2, \ldots, J$, and $n = 1, 2, \ldots, N$. The tensor $\mathcal{H}$ is then assembled by stacking along the third mode (front slices), resulting in $\mathcal{H}_{..n} = \mathbf{H}_\mathbf{Y}^{(n)}$. Using (3), $\mathcal{H}$ can be represented in scalar form as follows:

$$h_{i,j,n} = \sum_{r=1}^{R} m_{n,r} s_{n,i+j-1}. \tag{5}$$

Moreover, the $\mathcal{H}$ can be expressed as:

$$\mathcal{H} = \sum_{r=1}^{R} \mathbf{H}_\mathbf{S}^{(r)} \circ \mathbf{m}_r. \tag{6}$$

where $\mathbf{m}_r \in \mathbb{R}^K$ is the r-th column of $\mathbf{M}$ and $\mathbf{H}_\mathbf{S}^{(r)} \in \mathbb{R}^{I \times J}$ is a Hankel matrix constructed from the rows of $\mathbf{S}$, similarly to in (5).

Note that the tensor $\mathcal{H}$ in (6), created by the hankelization process, follows the BTD model in (1). In this way, after performing the BTD of $\mathcal{H}$, it is possible to estimate the sources from the matrices $\mathbf{H}_S^{(r)}$, for $r = 1, ..., R$.

As proposed in [29], Hankel-BTD can be applied to short-term ECG signals for AA extraction, specifically in records with persistent AF episodes of a single patient. The study reported better performance in BSS problems compared to matrix approaches. In the case of ECG signals, the source signals represent AA, VA, noise, breathing, and muscular activity, among others. In addition, the observed signals correspond to the ECG collected by the electrodes and the variable K represents the number of ECG leads.

IV. Database Description

Obtaining accurate estimates and conducting detailed analyzes are influenced by the data used during research, both in terms of quantity and quality. With large amounts of data, more robust patterns can be identified, leading to overall better results [23]. In addition, a broad set of databases is essential to recognize the complex nature of the subject under study. In this work, three databases were selected to assist in validating the suggested approach. The bases are as follows.

- **MIT-BIH Atrial Fibrillation (AFDB)** [24]: The database includes 25 recordings, mainly of paroxysmal AF. Records 00735 and 03665 were not included because they have heartbeats that were not audited, that is, they do not have the necessary information to extract the RRIs. Each recording is 10 hours long and contains two leads from the ECG, the signals are sampled at 250 Hz with 12-bit resolution in a range of ±10 mV.
- **MIT-BIH Long-Term Atrial Fibrillation (LTAFDB)** [24]: This database contains 84 long-term recordings from people who had persistent or paroxysmal AF. Each recording consists of two ECG leads, both digitized at a frequency of 128 Hz and 12-bit resolution, covering a range of ±20 mV. The duration of recordings can vary, but generally lasts between 24 and 25 hours.
- **MIT-BIH Normal Sinus Rhythm (NSRDB)** [24]: The database contains ECG records of patients with normal sinus rhythm, that is, without cardiac abnormalities. This database contains 18 ECG recordings of two leads, each recording lasting between 20 and 24 hours. The base sampling frequency is 128 Hz.

V. Methodology

Once the theoretical framework of the ECG signal mixing model was established, the next steps focus on preparing and processing the data, followed by the extraction of relevant features, which are essential for training and validating the proposed model.

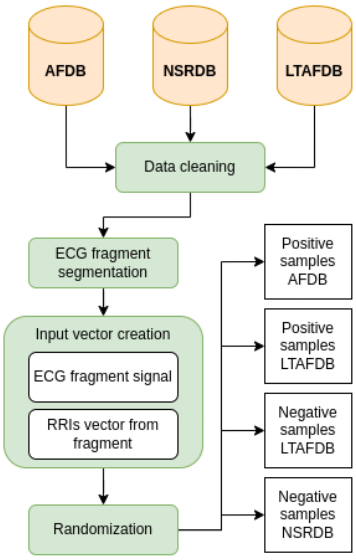


Fig. 2. Procedures for collecting and preparing input vectors.

TABLE I
TOTAL INPUT VECTORS EXTRACTED (BY DIAGNOSIS) FROM THE SELECTED DATABASES.

Database	Number of Segments (AF)	Number of Segments (NSR)	Total
AFDB	101497	-	101497
LTAFDB	618183	538877	1157060
NSRDB	-	361348	361348
Total	719680	900225	1619905

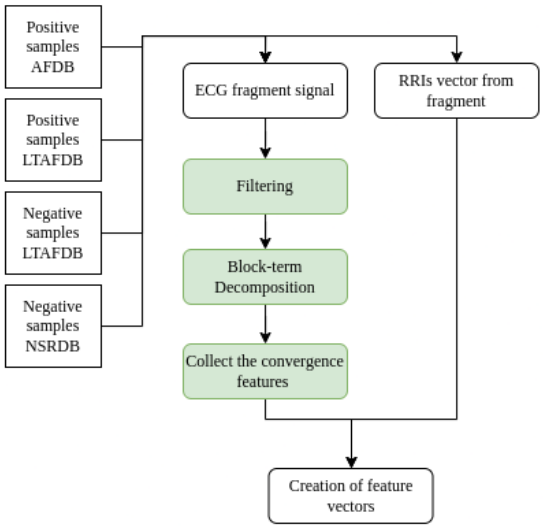


Fig. 3. Feature extraction flowchart.

A. Data Preparation

The data collection and preparation stage consists of a series of procedures that aim to ensure data consistency and increase variability between recordings for subsequent phases of the proposed technique. The central idea of this phase is to create an input vector for the feature extraction step of each segment of all recordings. In this context, the input vector refers to the data set formed by the ECG signal segment and its respective RRIs of the segment, which are collected from the recording header file. The flow chart of the preprocessing steps is shown in Fig. 2.

The process begins with the transfer of the selected databases to the local work environment, which already has recordings per patient separated in each *.dat* file. Then, the recordings are subjected to a cleaning stage, where those that do not have the audited signals are removed. Subsequently, excerpts of recordings that are not of interest to the research are discarded. For filtering, a type II bandm -pass filter was used, eliminating frequencies below 0.35 Hz and above 70 Hz. In this way, the sections diagnosed with AF or NSR are appropriately separated for the next step, which consists of fragmenting them into sections of 5 beats per segment (or 5 RRI).

The configuration adopted for the signal section size in this research is 5 RRIs. This choice is due to the attribute extraction methodology, which investigates features of different natures (BTD parameters and RRIs). In related work, these characteristics are approached with different segment sizes. RRIs are often explored in various ways, primarily in terms of histograms, which necessitate the use of longer signal segments. For example, in [25], a segmentation with 30 RRIs was proposed to create histograms. However, it becomes computationally impractical to calculate the estimated tensor on large signal segments using BTD. In studies related to BSS problems, which seek AA extraction through BTD, only a single heart beat (the QRS complex together with the TQ segment) is selected, as evidenced in the methodologies reported by [11], [21], [26]. In these works, it is also mentioned that the advantage of tensor methods compared to matrix-based methods, especially BTD, is their efficiency in extracting AA from short-term ECG recordings [11].

After creating the data packages (input vectors) formed by the combination of the ECG signal and their respective RRIs, a randomization step was applied with the aim of increasing generalization of the model and avoiding overfitting. Tab. I presents a quantitative analysis of ECG segments extracted and classified by type of cardiac rhythm (AF and NSR) in the selected databases (AFDB, LTAFDB and NSRDB).

B. Feature Extraction

The next steps are related to the computation of the feature vectors from the input vectors collected in the previous step. Fig. 3 summarizes the feature extraction procedure using a flow diagram. The main idea of the proposed methodology is to use the BTD convergence parameters, along with the RRI, as input features for the classifiers.

The motivation behind the use of the BTD convergence parameters as discriminant features is the fact that the Hankel-BTD of an AF segment behaves differently for AF and NSR. In fact, the Hankel-BTD of a healthy heart rate segment converges less efficiently than in a segment with AF. This difference would be reflected in the algorithm's convergence and optimization indicators, enabling effective separation.

This is due to the fact that Hankel-BTD with ECG signals is only guaranteed to be unique when applied to signals with the presence of AF. In fact, when the ECG signal has AF, the atrial sources can be modeled by all-pole models, also known as an exponential model [20], as follows:

$$s_{r,n} = \sum_{l=1}^{L_r} \lambda_{l,r} z_{l,r}^{(n-1)} \tag{7}$$

for $r = 1, \ldots, R$, where L_r is the number of exponential terms, $z_{l,r}$ is the l-th pole of the r-th source and $\lambda_{l,r}$ is the scaling coefficient.

This modeling of the atrial sources ensures that the uniqueness conditions of the Hankel-BTD are satisfied [28]. In fact, it can be demonstrated that, when atrial sources are approximated by single-pole models, the Hankel matrices used in the Hankel-BTD method can be expressed through the Vandermonde decomposition, in the following way:

$$\mathbf{H}_{\mathbf{S}}^{(r)} = V_r \text{diag}(\lambda_{1,r}, \lambda_{2,r}, ..., \lambda_{L_r,r}) \hat{V}_r^T, \tag{8}$$

with

$$V_r = \begin{bmatrix} 1 & 1 & \ldots & 1 \\ z_{1,r} & z_{2,r} & \ldots & z_{L_r,r} \\ \vdots & \vdots & \ddots & \vdots \\ z_{1,r}^{I-1} & z_{2,r}^{I-1} & \ldots & z_{L_r,r}^{I-1} \end{bmatrix} \in \mathbb{R}^{I \times L_r}, \tag{9}$$

e

$$\hat{V}_r = \begin{bmatrix} 1 & 1 & \ldots & 1 \\ z_{1,r} & z_{2,r} & \ldots & z_{L_r,r} \\ \vdots & \vdots & \ddots & \vdots \\ z_{J,r}^{J-1} & z_{2,r}^{J-1} & \ldots & z_{L_r,r}^{J-1} \end{bmatrix} \in \mathbb{R}^{J \times L_r}. \tag{10}$$

For the case where the poles are distinct, as noted in [20], the Vandermonde matrix, whose number of columns is determined by $L_r \leq I, J$, exhibits the full-column rank. Consequently, if the matrix M does not contain proportional columns, the BTD in (6) is essentially unique. In instances where the poles are equal, the uniqueness can still be ensured under less restrictive conditions.

The experimental results of previous work corroborate the efficacy of the Hankel-BTD approach when dealing with AF ECG signals [21], [22]. However, when estimating Hankel-BTD in a signal segment that does not have AF, there is no guarantee of the uniqueness of the decomposition. In this case, it is expected that the parameters that evaluate the quality and speed of BTD convergence will differ significantly from the case of segments with AF. Due to this, this study proposes that the convergence parameters of the Hankel-BTD should be used as discriminative features between NSR and AF. In particular, the following parameters of the *ll1_rnd* function (from Tensorlab [48]) are considered:

- **relfval**: The difference in the value of the objective function between two successive iterations, in relation to its initial value;
- **relerr**: The relative error between the original tensor and its estimated BTD model.
- **fval**: The value of the objective function in each iteration;

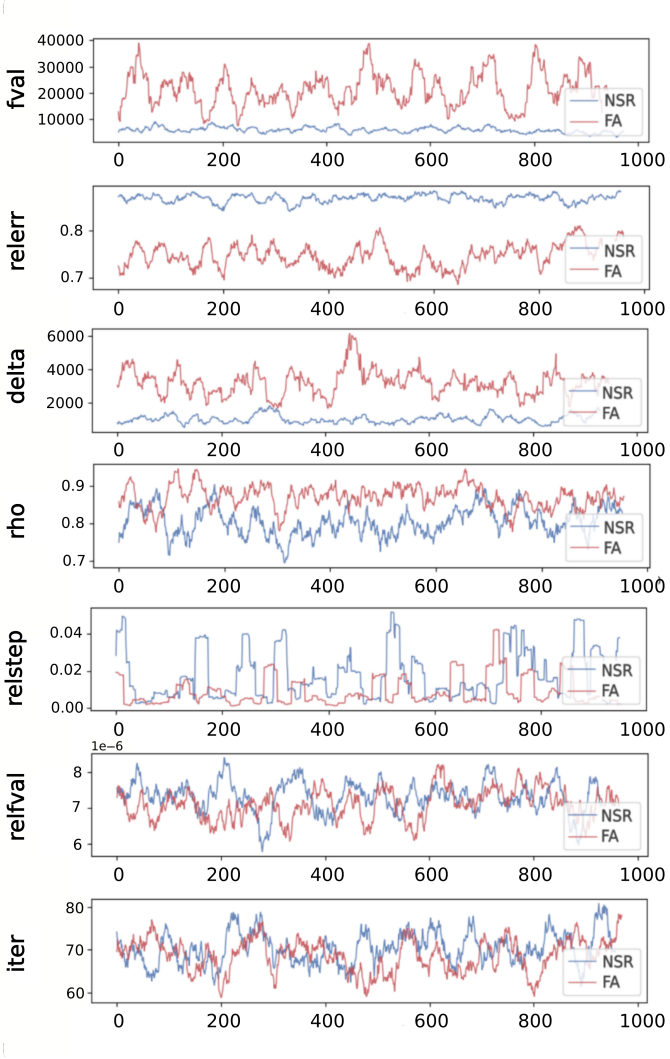


Fig. 4. Samples of the Hankel-BTD convergence parameters, for both classes (NSR and AF).

- **relstep**: The step size relative to the norm of the current iteration in each iteration;
- **delta**: The radius of the confidence region in each iteration;
- **rho**: The reliability in each iteration attempt;
- **iter**: Number of iterations until **fval**;

Fig. 4 shows 1000 samples of the above convergence parameters, for both classes (NSR and AF). The following steps are considered in Fig. 4:

- **Sample Selection**: 1000 samples were randomly selected from the dataset.
- **Z-score Filter**: The limit established for filtering was 3.
- **Convolution**: The final result was generated through a convolution window of size 25 among the selected samples.
- **Signal average**: This procedure was repeated 1000 times, and then the average of the executions was calculated to plot the graph.

It is possible to see from this figure that these parameters have a good ability to discriminate the two classes. The

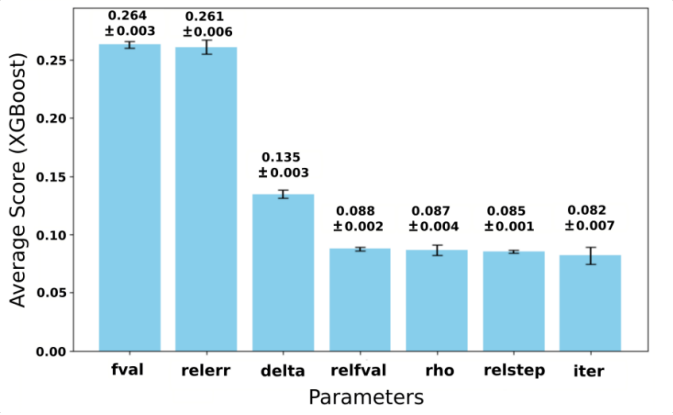


Fig. 5. Relevance scores in XGBoost of the Hankel-BTD convergence parameters.

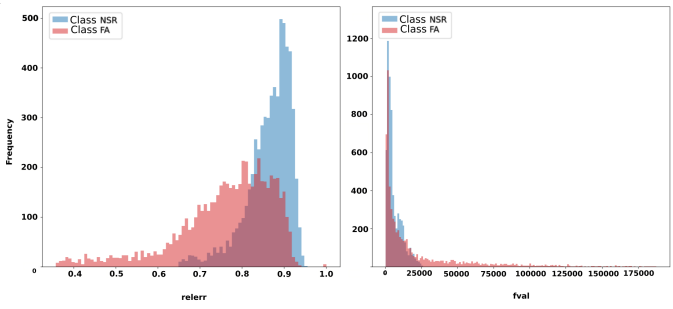


Fig. 6. Histograms of *relerr* and *fval* for the AF and NSR classes.

discriminatory ability of these parameters can also be verified in Fig. 5, which shows the relevance score obtained using the XGBoost algorithm. The relevance of the features is calculated as the total normalized reduction of the criterion driven by a given feature, and is also known as the Gini index. These scores provide an understanding of the influence of each characteristic on the decision-making process of the generated model. Generally, features with higher importance scores are considered more relevant for classification. These results were obtained from all collected samples and using a cross validation with 10 folds. After processing each fold, the mean and standard deviation of the relevance scores were calculated and highlighted just above each bar in Fig. 5.

Fig. 6 shows the histograms of the parameters *relerr* and *fval*, for the two classes. These two parameters provided the highest values of the relevance scores in Fig. 5. It is observed that their histograms do not have a significant overlap, suggesting that the AF and NSR classes can be easily discriminated based on these characteristics if used together.

It can be observed in Figs. 4 and 5 that the three parameters that contribute the most to discriminate the AF and NSR classes are *relerr*, *fval* and *delta*. Based on these results presented, these three parameters will be considered in the present work. On the other hand, the indicators that do not contribute significantly and, therefore, will not be used later are: *relfval*, *rho*, *relstep* and *iter*.

The final attributes used by the proposed methodology are these three BTD parameters (*relerr*, *fval* and *delta*), along with the RRI values. These two types of attributes (convergence parameters of BTD and RRIs), when applied together, can provide efficient discriminatory information for the detection of AF.

TABLE II
DISTRIBUTION OF FEATURE VECTORS (SEPARATED BY RECORDING).

Database	Label	Recording	Quantity	Total
AFDB	AF	04126, 04746, 04908, 04936, 05121, 06426, 06995, 07162, 07859, 07879, 07910, 08215, 08219, 08378, 08405, 08434 e 08455	200	3400
		04043	1000	1000
		05261	178	178
		04048	158	158
		04015	99	99
		06453	84	84
		05091	22	22
NSRDB	NSR	16265, 16272, 16273, 16420, 16483, 16539, 16773, 16786, 16795, 17052, 17453, 18177, 18184, 19088, 19090, 19093, 19140 e 19830	200	3600
LTAFDB	NSR	00, 01, 03, 05, 06, 07, 08, 10, 100, 101, 102, 103, 104 e 105	200	2800
LTAFDB	AF	00, 01, 03, 06, 07, 100, 101, 102, 103 e 104	200	2000
		08	34	34
		05	11	11
Total				**13386**

Due to the quantity and the average time required to perform the complete extraction of a sample, the feature extraction process used 4941 samples for the AFDB database with the AF label, 3600 samples for the NSRDB database for the NSR label, and finally, 2800 and 2000 samples for the NSR and AF labels in the LTAFDB database. In total, 13341 samples were obtained, which will serve as input data for the classifiers. Tab. II presents the distribution of feature vectors ready for validation, which are separated by recording from the LTAFDB, NSRDB and AFDB databases.

C. Classification Algorithms

Tree-based algorithms are popular machine learning methods that are frequently used to handle complex, high-dimensional datasets because of their capacity to model nonlinear relationships. These models use hierarchical data partitioning, which enables them to adjust to diverse feature patterns, a characteristic that is especially beneficial in high-dimensional settings, such as those in medical diagnostics. Many works in the literature, such as [40] and [41], demonstrate that tree-based models consistently achieve greater accuracy when applied to complex datasets. As a result, they are often the preferred choice for applications that require a

TABLE III
RESULTS OBTAINED WITH ADJUSTED HYPERPARAMETERS AND K-FOLD.

	Gradient Boosting		
Fold	**ACC** (%)	**SEN** (%)	**SPE** (%)
3	97.16	97.68	96.56
Average	95.97 ± 0.63	**96.70** ± 0.68	95.14 ± 0.97
	Random Forest		
Fold	**ACC** (%)	**SEN** (%)	**SPE** (%)
4	96.87	98.04	95.56
Average	95.72 ± 0.86	**96.89** ± 0.79	94.40 ± 1.22
	Extremely Random Trees		
Fold	**ACC** (%)	**SEN** (%)	**SPE** (%)
9	97.82	98.18	97.43
Average	96.66 ± 0.60	**97.42** ± 0.46	95.84 ± 1.28
	XGBoost		
Fold	**ACC** (%)	**SEN** (%)	**SPE** (%)
10	97.34	98.57	95.95
Average	97.21 ± 0.42	**97.81** ± 0.58	96.54 ± 0.63
	LightGBM		
Fold	**ACC** (%)	**SEN** (%)	**SPE** (%) (%)
4	98.01	98.92	96.99
Average	97.48 ± 0.43	**97.86** ± 0.74	97.05 ± 0.52

subtle interaction of characteristics, including ECG-based AF detection [49]. Due to these reasons, this study made use of the following tree-based classifiers: Random Forest, Extremely Random Trees, Gradient Boosting, Extreme Gradient Boosting, and the LightGBM.

VI. RESULTS AND DISCUSSIONS

This section presents the simulation results obtained with the aim of evaluating the performance and behavior of the proposed technique. The hyperparameter was tuned using the grid search technique. Although computationally expensive, the grid search technique guaranteed more reliable results for AF detection. Furthermore, each subsection provides both quantitative results and interpretative insights to support the clinical relevance of our findings.

A. Results with cross validation (k-fold)

The results obtained by each classifier using the set of optimal hyperparameters are presented in Tab. III. In the table, folds achieving the highest sensitivity are highlighted, and mean ± standard deviation are reported for accuracy (ACC), sensitivity (SEN), and specificity (SPE). ACC measures the overall proportion of correctly classified samples; SEN quantifies correctly detected AF cases, helping to reduce false negatives; SPE reflects correctly classified NSR cases. Sensitivity was selected for fold evaluation due to its clinical relevance in handling class imbalance.

The data used for this procedure were previously selected from a portion of 80% of the total set, while the remaining 20% was kept separated as a test set, that is, a portion of the data that will be used only to evaluate the performance of the model. Using stratified 10-fold cross-validation with exhaustive grid search, all tree-based classifiers achieved high performance (mean ACC $\geq$ 95%, SEN $\geq$ 96%). LightGBM led with an average sensitivity of 97.86% ± 0.74%, demonstrating robust AF detection. Specificity trailed sensitivity, reflecting a focus on minimizing missed AF cases in imbalanced data. This analysis highlights the model's ability to consistently detect AF across folds and justifies prioritizing sensitivity in clinical screening scenarios.

TABLE IV
RESULTS WITH TEST DATA.

Classifier	**ACC (%)**	**SEN (%)**	**SPE (%)**
LightGBM	97.04	97.44	96.59
XGBoost	96.63	96.93	96.28
Extremely Random Trees	95.90	97.07	94.59
Gradient Boosting	95.11	95.30	94.89
Random Forest	94.65	95.66	93.51

TABLE V
CONFUSION MATRIX FOR THE BEST CASE.

	Positive Predicted	**Negative Predicted**
Positive Real	1368	36
Negative Real	44	1191

B. Results with test data

The results presented in this subsection use all the samples from the previous section to train the model, totaling 10548 samples for training. The created models will be validated using test data (20%), representing 2637 samples. This portion of the data has not been used for training or testing in any previous simulations in this research. Tab. IV presents the performance of the different algorithms based on the hold-out validation technique in terms of accuracy, sensitivity, and specificity, respectively.

LightGBM again topped all measures (ACC 97.04%, SEN 97.44%, SPE 96.59%), while Random Forest showed the largest drop (ACC 94.65%, SEN 95.66%). This consistency validates the generalizability of LightGBM as the final model. These hold-out results affirm the model's readiness for deployment, demonstrating stable performance on fully unseen data.

Therefore, this research considers the results presented in Tab. IV for the LightGBM classifier as the conclusive result of the proposed technique. The hyperparameters used by the LightGBM in this best case are the following:

- num_leaves: 50
- learning_rate: 0.1
- max_depth: *None*
- n_estimators: 300

As mentioned above, these parameters were selected by an exhaustive search using the *grid search* method.

The confusion matrix corresponding to this best case is presented in Tab. V. It can be seen in this table that the difference in the number of correctly labeled classes stands out, with 1368 (VP) and 1191 (VN), due to the imbalance between the classes in the dataset used. The total number of samples tested was 2559 (100%), with 1404 (53.2%) labeled AF and 1235 (46.8%) as NSR. The results demonstrated the model's superior performance in identifying positive samples, once again minimizing critical classification errors. This is evidenced by the low incidence of false negatives (36 episodes),

TABLE VI
METRICS OBTAINED FROM THE 95% CONFIDENCE INTERVAL.

Metric	Inferior Limit 95%	Superior Limit 95%
ACC (%)	97.03	97.07
SEN (%)	97.48	97.54
SPE (%)	96.51	96.57

although the number of positively labeled samples is slightly higher (3.2%).

In our study, the classifiers achieved an AUC of 0.99, demonstrating high performance in the binary classification task that distinguishes between the AF and NSR classes. The Area Under the Curve (AUC) refers to the area under the Receiver Operating Characteristic (ROC) curve and serves as a comprehensive measure of the discriminative capacity of a model. An AUC value closer to 1.0 indicates superior classification performance, whereas a value near 0.5 suggests a performance equivalent to random chance. In contrast, baseline classifiers used for comparative purposes yielded less favorable results. For example, logistic regression achieved an accuracy of 76.68% and an AUC of 0.87, while the Gaussian Naive Bayes classifier attained an accuracy of 82.18% with an AUC of 0.88. Lastly, the algorithm for the nearest neighbors k reached an accuracy of 85.59% and an AUC of 0.92.

C. Confidence interval

The confidence intervals were also calculated, which corresponds to a range of values that indicates the probability that a population parameter is contained within this interval. To perform the calculation, the sampling method called bootstrap was used, which consists of obtaining random samples from the original sample. After 1000 simulations, which involved bootstrap sampling in each iteration, the lower and upper limits of the confidence interval 95% were obtained for each evaluation metric, shown in Tab. VI.

Bootstrap analysis yielded narrow 95% intervals for ACC (97.03–97.07%), SEN (97.48–97.54%), and SPE (96.51–96.57%), reflecting precise, reliable performance estimates. The tight confidence bounds underscore the robustness of our primary metrics, reinforcing confidence in clinical applicability.

D. Comparison with related works

The Tab. VII summarizes the results obtained in works related to this research. The study to be used as a comparative reference is presented in [25], as it uses the same databases. This table provides an overview of various studies on AF detection, highlighting datasets such as AFDB, NSRDB, and LTAFDB, along with performance metrics such as accuracy, sensitivity, and specificity. Most studies mainly use AFDB, with some incorporating datasets such as ADB. Notable is that some work includes [45], achieving the highest specificity (99.63%), and [44], which excels in all metrics. The proposed approach combines multiple datasets, offering balanced results (97.04% ACC, 97.44% SEN, 96.59% SPE), showcasing its reliability over methods focused on single datasets.

TABLE VII
RESULTS OF RELATED WORKS (IN CRESCENT CHRONOLOGICAL ORDER).

	AFDB	NSRDB	LTAFDB	ADB	ACC (%)	SEN (%)	SPE (%)
[42]	✓				94.75	91.38	93.5
[12]	✓				-	94.4	97.2
[43]				✓	93.7	95.1	92.0
[44]	✓				98.6	98.9	98.4
[45]	✓				98.9	97.93	99.63
[46]	✓			✓	-	100	97.6
[47]	✓				97.6	98.0	97.4
[25]	✓	✓	✓		96.97	95.24	99.94
Our	✓	✓	✓		97.04	97.44	96.59

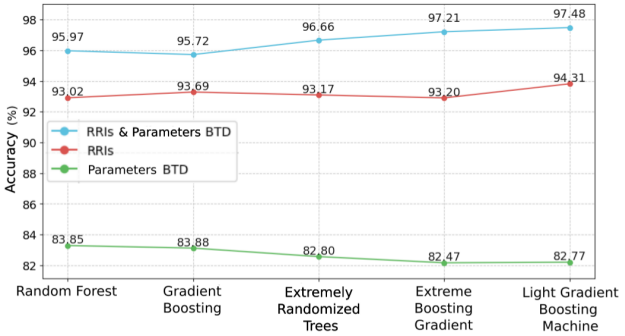


Fig. 7. Comparison between characteristic configurations.

These results also highlight improvements in accuracy and sensitivity compared to the state-of-the-art approach by [25]. Furthermore, while [25] uses a 30 RRI ECG signal segment to create feature vectors, the present work reduces this to a 5 RRI segment and a reduction of 16. 67% showing efficiency with short ECG segments in time.

E. Contribution of BTD parameters

Integrating tensor-decomposition convergence parameters with RRI features improved precision across all classifiers, most notably a 3.17% gain for LightGBM. These BTD metrics capture the nuances of the ECG signal beyond RRIs, confirming their value in enhancing the detection of AF. This finding validates the proposed feature engineering strategy and suggests potential for broader application in arrhythmia classification.

It can be seen from this figure that the approach using only RRIs already has relevance in detecting AF, which is also reported in related works. However, when the RRIs were combined with the BTD parameters, there was an improvement in the accuracy performance in all techniques. For LightGBM, there is a significant increase of 3.17% in accuracy. This result highlights the relevance and positive impact of the combination of RRI and BTD parameters as features for the detection of AF.

VII. CONCLUSION

This work contributed to the detection of AF in cardiac monitoring using ECG recordings. Key achievements include

the preparation of a segmented, preprocessed ECG signal repository and the reproduction of RRI extraction, widely used for the detection of arrhythmias. Furthermore, the research proposed a tensor-based technique for AF detection, using parameters that evaluate the convergence of the BTD. As Hankel-BTD convergence behaves differently in the case of AF and NSR ECG, some parameters obtained during the calculation of the BTD contain significant discriminant information for the detection of AF. Tree-based classifiers validated the approach, with hyperparameter adjustments yielding significant performance improvements.

The findings demonstrated the effectiveness of the method, achieving robust performance metrics, including average accuracy, sensitivity, and specificity above 97%, and validated the relevance of the BTD parameters for the detection of AF in conjunction with RRIs.

Future work could include validating this approach on 12-lead ECG databases, improving preprocessing techniques, leveraging Optuna for hyperparameter tuning, and expanding datasets to improve generalizability. Deep learning and transfer learning approaches are also promising avenues for further exploration.

REFERENCES

[1] A. G. Garwal, "Mais de 289 mil pessoas morreram de doenças cardiovasculares em 2019," Agência Brasil, 2019. [Online]. Available: http://agenciabrasil.ebc.com.br/saude/noticia/2019-09/mais-de-289-mil-pessoas-morreram-de-doencas-cardiovasculares-em-2019. [Accessed: May 19, 2024].

[2] G. Lip, L. Fauchier, B. Freedman, I. Gelder, A. Natale, C. Gianni, S. Nattel, T. Potpara, M. Rienstra, H. Tse, and D. Lane, "Atrial fibrillation," Nature Reviews Disease Primers, vol. 2, 2016. doi: 10.1038/nrdp.2016.16.

[3] C. Morillo, A. Banerjee, P. Perel, D. A. Wood, and X. Jouven, "Atrial fibrillation: the current epidemic," Journal of Geriatric Cardiology: JGC, vol. 14, pp. 195-203, 2017. doi: 10.11909/j.issn.1671-5411.2017.03.011.

[4] L. Mainardi, L. Sörnmo, and S. Cerutti, "Understanding atrial fibrillation: The signal processing contribution, part II," Synthesis Lectures on Biomedical Engineering, Morgan & Claypool Publishers LLC, vol. 3, no. 1, pp. 1-139, Jan. 2008. doi: 10.2200/s00153ed1v01y200809bme025.

[5] L. Kjerpeseth, J. Igland, R. Selmer, H. Ellekjær, A. Tveit, T. Berge, S. Kalstø, I. Christophersen, M. Myrstad, E. Skovlund, G. Egeland, G. Tell, and I. Ariansen, "Prevalence and incidence rates of atrial fibrillation in Norway 2004–2014," Heart, vol. 107, pp. 201-207, 2020. doi: 10.1136/heartjnl-2020-316624.

[6] J. Delaney, X. Yin, J. Fontes, E. Wallace, A. Skinner, N. Wang, B. Hammill, E. Benjamin, L. Curtis, and S. Heckbert, "Hospital and clinical care costs associated with atrial fibrillation for Medicare beneficiaries in the cardiovascular health study and the Framingham heart study," Sage Open Medicine, vol. 6, 2018. doi: 10.1177/2050312118759444.

[7] J. Mant, D. A. Fitzmaurice, F. D. R. Hobbs, S. Jowett, E. T. Murray, R. Holder, M. Davies, and G. Y. H. Lip, "Accuracy of diagnosing atrial fibrillation on electrocardiogram by primary care practitioners and interpretative diagnostic software: analysis of data from screening for atrial fibrillation in the elderly (SAFE) trial," BMJ, vol. 335, no. 7616, p. 380, Jun. 2007. doi: 10.1136/bmj.39227.551713.ae.

[8] A. Rizwan, A. Zoha, I. B. Mabrouk, H. M. Sabbour, A. S. Al-Sumaiti, A. Alomainy, M. A. Imran, and Q. H. Abbasi, "A review on the state of the art in atrial fibrillation detection enabled by machine learning," IEEE Reviews in Biomedical Engineering, vol. 14, pp. 219-239, 2021. doi: 10.1109/RBME.2020.2976507.

[9] P. M. R. de Oliveira and V. Zarzoso, "Source analysis and selection using block term decomposition in atrial fibrillation," in Latent Variable Analysis and Signal Separation, pp. 46-56, 2018. doi: 10.1007/978-3-319-93764-9_5.

[10] P. M. R. de Oliveira and V. Zarzoso, "Temporal stability of block term decomposition in noninvasive atrial fibrillation analysis," in Proc. 2018 52nd Asilomar Conference on Signals, Systems, and Computers, 2018. doi: 10.1109/ACSSC.2018.8645164.

[11] P. M. R. de Oliveira, V. Zarzoso, and C. A. R. Fernandes, "Coupled tensor model of atrial fibrillation ECG," in Proc. 2020 28th European Signal Processing Conference (EUSIPCO), IEEE, 2021. doi: 10.23919/Eusipco47968.2020.9287494.

[12] K. Tateno and L. Glass, "Automatic detection of atrial fibrillation using the coefficient of variation and density histograms of RR and RR intervals," Medical and Biological Engineering and Computing, vol. 39, no. 6, pp. 664-671, Nov. 2001. doi: 10.1007/BF02345439.

[13] F. L. Hitchcock, "The expression of a tensor or a polyadic as a sum of products," Journal of Mathematics and Physics, Wiley, vol. 6, no. 1-4, pp. 164-189, Apr. 1927. doi: 10.1002/sapm192761164.

[14] L. De Lathauwer, "Decompositions of a higher-order tensor in block terms—part II: Definitions and uniqueness," SIAM Journal on Matrix Analysis and Applications, vol. 30, no. 3, pp. 1033-1066, Jan. 2008. doi: 10.1137/070690729.

[15] A. Belouchrani, K. Abed-Meraim, J. Cardoso, and A. Moulines, "A blind source separation technique using second-order statistics," IEEE Transactions on Signal Processing, vol. 45, pp. 434-444, 1997. doi: 10.1109/78.554307

[16] J. Rieta, F. Castells, C. Sanchez, V. Zarzoso, and J. Millet, "Atrial activity extraction for atrial fibrillation analysis using blind source separation," IEEE Transactions on Biomedical Engineering, vol. 51, no. 7, pp. 1176-1186, Jul. 2004. doi: 10.1109/tbme.2004.827272.

[17] O. Debals and L. De Lathauwer, "Stochastic and deterministic tensorization for blind signal separation," in Lecture Notes in Computer Science, Springer International Publishing, 2015, pp. 3-13. ISBN 9783319224824. doi: 10.1007/978-3-319-22482-4_1.

[18] L. De Lathauwer, "Blind separation of exponential polynomials and the decomposition of a tensor in rank-(lr, lr, 1) terms," SIAM Journal on Matrix Analysis and Applications, vol. 32, no. 4, pp. 1451-1474, Oct. 2011. doi: 10.1137/100805510.

[19] L. N. Ribeiro, A. R. Hidalgo-Munoz, and V. Zarzoso, "Atrial signal extraction in atrial fibrillation electrocardiograms using a tensor decomposition approach," in Proc. 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2015. doi: 10.1109/EMBC.2015.7320000.

[20] V. Zarzoso, "Parameter estimation in block term decomposition for noninvasive atrial fibrillation analysis," in 2017 IEEE 7th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), IEEE, 2017. doi: 10.1109/CAMSAP.2017.8313173.

[21] P. M. R. de Oliveira and V. Zarzoso, "Temporal stability of block term decomposition in noninvasive atrial fibrillation analysis," in 2018 52nd Asilomar Conference on Signals, Systems, and Computers, IEEE, 2018. doi: 10.1109/ACSSC.2018.8645164.

[22] P. M. R. de Oliveira and V. Zarzoso, "Löwner-based tensor decomposition for blind source separation in atrial fibrillation ECGs," in 2019 27th European Signal Processing Conference (EUSIPCO), IEEE, 2019. doi: 10.23919/EUSIPCO.2019.8902493.

[23] A. Halevy, P. Norvig, and F. Pereira, "The unreasonable effectiveness of data," IEEE Intelligent Systems, vol. 24, no. 2, pp. 8-12, Mar. 2009. [Online]. doi: 10.1109/mis.2009.36.

[24] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," Circulation, vol. 101, no. 23, pp. e215-e220, 2000. doi: 10.1161/01.CIR.101.23.e215.

[25] J. Duan, Q. Wang, B. Zhang, C. Liu, C. Li, and L. Wang, "Accurate detection of atrial fibrillation events with R-R intervals from ECG signals," PLOS ONE, vol. 17, no. 8, p. e0271596, Aug. 2022. doi: 10.1371/journal.pone.0271596.1.

[26] P. M. R. de Oliveira and V. Zarzoso, "Source analysis and selection using block term decomposition in atrial fibrillation," in Lecture Notes in Computer Science, Springer International Publishing, 2018, pp. 46-56. ISBN 9783319937649. doi: 10.1007/978-3-319-93764-9_5.

[27] Rieta, J. J., Castells, F., Sánchez, C., Zarzoso, V., & Millet, J. (2004). Atrial activity extraction for atrial fibrillation analysis using blind source separation. IEEE Transactions on Bio-Medical Engineering, 51(7), 1176–1186. doi:10.1109/TBME.2004.827272.

[28] De Lathauwer, L. (2011). Blind separation of exponential polynomials and the decomposition of a tensor in rank-(L_r,L_r,1) terms. SIAM Journal on Matrix Analysis and Applications: A Publication of the Society for Industrial and Applied Mathematics, 32(4), 1451–1474. doi:10.1137/100805510.

[29] Ribeiro, L. N., Hidalgo-Munoz, A. R., & Zarzoso, V. (2015, August). Atrial signal extraction in atrial fibrillation electrocardiograms using a tensor decomposition approach. 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).

Presented at the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan. doi:10.1109/embc.2015.7320000

[30] Goulart, J. H. de M., de Oliveira, P. M. R., Farias, R. C., Zarzoso, V., & Comon, P. (2020). Alternating group lasso for block-term tensor decomposition and application to ECG source separation. IEEE Transactions on Signal Processing: A Publication of the IEEE Signal Processing Society, 68, 2682–2696. doi:10.1109/tsp.2020.2985591

[31] de Oliveira, P. M. R., Goulart, J. H. de M., Fernandes, C. A. R., & Zarzoso, V. (2022). Blind source separation in persistent atrial fibrillation electrocardiograms using block-term tensor decomposition with löwner constraints. IEEE Journal of Biomedical and Health Informatics, 26(4), 1538–1548. doi:10.1109/JBHI.2021.3108699

[32] Andersen, R. S., Peimankar, A., & Puthusserypady, S. (2019). A deep learning approach for real-time detection of atrial fibrillation. Expert Systems with Applications, 115, 465–473. doi:10.1016/j.eswa.2018.08.011

[33] Siontis, K. C., Noseworthy, P. A., Attia, Z. I., & Friedman, P. A. (2021). Artificial intelligence-enhanced electrocardiography in cardiovascular disease management. Nature Reviews. Cardiology, 18(7), 465–478. doi:10.1038/s41569-020-00503-2

[34] Khurshid, S., Friedman, S., Reeder, C., Di Achille, P., Diamant, N., Singh, P., Lubitz, S. A. (2022). ECG-based deep learning and clinical risk factors to predict Atrial Fibrillation. Circulation, 145(2), 122–133. doi:10.1161/CIRCULATIONAHA.121.057480

[35] Van Gelder, I. C., Rienstra, M., Bunting, K. V., Casado-Arroyo, R., Caso, V., Crijns, H. J. G. M., … ESC Scientific Document Group. (2024). 2024 ESC Guidelines for the management of atrial fibrillation developed in collaboration with the European Association for Cardio-Thoracic Surgery (EACTS). European Heart Journal, 45(36), 3314–3414. doi:10.1093/eurheartj/ehae176

[36] Roth, G. A., Mensah, G. A., Johnson, C. O., Addolorato, G., Ammirati, E., Baddour, L. M., GBD-NHLBI-JACC Global Burden of Cardiovascular Diseases Writing Group. (2020). Global Burden of cardiovascular diseases and risk factors, 1990-2019: Update from the GBD 2019 Study. Journal of the American College of Cardiology, 76(25), 2982–3021. doi:10.1016/j.jacc.2020.11.010

[37] Chugh, S. S., Havmoeller, R., Narayanan, K., Singh, D., Rienstra, M., Benjamin, E. J., Murray, C. J. L. (2014). Worldwide epidemiology of atrial fibrillation: a Global Burden of Disease 2010 Study. Circulation, 129(8), 837–847. doi:10.1161/CIRCULATIONAHA.113.005119

[38] Turakhia, M. P., Guo, J. D., Keshishian, A., Delinger, R., Sun, X., Ferri, M., Hlavacek, P. (2023). Contemporary prevalence estimates of undiagnosed and diagnosed atrial fibrillation in the United States. Clinical Cardiology, 46(5), 484–493. doi:10.1002/clc.23983

[39] Elliott, A. D., Middeldorp, M. E., Van Gelder, I. C., Albert, C. M., & Sanders, P. (2023). Author Correction: Epidemiology and modifiable risk factors for atrial fibrillation. Nature Reviews. Cardiology, 20(6), 429. doi:10.1038/s41569-023-00834-w

[40] Grinsztajn, L., Oyallon, E., & Varoquaux, G. (2022). Why do tree-based models still outperform deep learning on tabular data? doi:10.48550/ARXIV.2207.08815

[41] Uddin, S., & Lu, H. (2024). Confirming the statistically significant superiority of tree-based machine learning algorithms over their counterparts for tabular data. PloS One, 19(4), e0301541. doi:10.1371/journal.pone.0301541

[42] Young, B., Brodnick, D., & Spaulding, R. (2003). A comparative study of a hidden Markov model detector for atrial fibrillation. Neural Networks for Signal Processing IX: Proceedings of the 1999 IEEE Signal Processing Society Workshop (Cat. No.98TH8468). Presented at the Neural Networks for Signal Processing IX: 1999 IEEE Signal Processing Society Workshop, Madison, WI, USA. doi:10.1109/nnsp.1999.788166

[43] Li, Y., Tang, X., Wang, A., & Tang, H. (2017). Probability density distribution of delta RR intervals: a novel method for the detection of atrial fibrillation. Australasian Physical & Engineering Sciences in Medicine, 40(3), 707–716. doi:10.1007/s13246-017-0554-2

[44] Czabanski, R., Horoba, K., Wrobel, J., Matonia, A., Martinek, R., Kupka, T., Leski, J. M. (2020). Detection of atrial fibrillation episodes in long-term heart rhythm signals using a support vector machine. Sensors (Basel, Switzerland), 20(3), 765. doi:10.3390/s20030765.

[45] Hu, Y., Zhao, Y., Liu, J., Pang, J., Zhang, C., & Li, P. (2020). An effective frequency-domain feature of atrial fibrillation based on time-frequency analysis. BMC Medical Informatics and Decision Making, 20(1), 308. doi:10.1186/s12911-020-01337-1

[46] Lown, M., Brown, M., Brown, C., Yue, A. M., Shah, B. N., Corbett, S. J., Little, P. (2020). Machine learning detection of Atrial Fibrillation using wearable technology. PloS One, 15(1), e0227401. doi:10.1371/journal.pone.0227401

[47] Hirsch, G., Jensen, S. H., Poulsen, E. S., & Puthusserypady, S. (2021). Atrial fibrillation detection using heart rate variability and atrial activity: A hybrid approach. Expert Systems with Applications, 169(114452), 114452. doi:10.1016/j.eswa.2020.114452

[48] Vervliet, N., Debals, O., & De Lathauwer, L. (2016, November). Tensorlab 3.0 — Numerical optimization strategies for large-scale constrained and coupled matrix/tensor factorization. 2016 50th Asilomar Conference on Signals, Systems and Computers. Presented at the 2016 50th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA. doi:10.1109/acssc.2016.7869679

[49] Shao, M., Bin, G., Wu, S., Bin, G., Huang, J., & Zhou, Z. (2018). Detection of atrial fibrillation from ECG recordings using decision tree ensemble with multi-level features. Physiological Measurement, 39(9), 094008. doi:10.1088/1361-6579/aadf48

[50] Houben, R. P. M., de Groot, N. M. S., & Allessie, M. A. (2010). Analysis of fractionated atrial fibrillation electrograms by wavelet decomposition. IEEE Transactions on Biomedical Engineering, 57(6), 1388–1398. https://doi.org/10.1109/tbme.2009.2037974

[51] Passman, R. (2021). Mobile health technologies in the diagnosis and management of atrial fibrillation. Current Opinion in Cardiology, 37(1), 1–9. https://doi.org/10.1097/hco.0000000000000930

[52] Zhang, X., Li, J., Cai, Z., Zhang, L., Chen, Z., & Liu, C. (2021). Over-fitting suppression training strategies for deep learning-based atrial fibrillation detection. Medical & Biological Engineering & Computing, 59(1), 165–173. https://doi.org/10.1007/s11517-020-02292-9

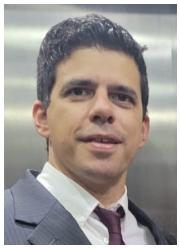
[53] Cheng, Y., Hu, Y., Hou, M., Pan, T., He, W., & Ye, Y. (2020). Atrial fibrillation detection directly from compressed ECG with the prior of measurement matrix. Information, 11(9), 436. https://doi.org/10.3390/info11090436

[54] Liaqat, S., Dashtipour, K., Zahid, A., Assaleh, K., Arshad, K., & Ramzan, N. (2020). Detection of atrial fibrillation using a machine learning approach. Information, 11(12), 549. https://doi.org/10.3390/info11120549

[55] Fujita, H., & Cimr, D. (2019). Computer aided detection for fibrillations and flutters using deep convolutional neural network. Information Sciences, 486, 231–239. https://doi.org/10.1016/j.ins.2019.02.065

[56] Acharya, U. R., Fujita, H., Lih, O. S., Hagiwara, Y., Tan, J. H., & Adam, M. (2017). Automated detection of arrhythmias using different intervals of tachycardia ECG segments with convolutional neural network. Information Sciences, 405, 81–90. https://doi.org/10.1016/j.ins.2017.04.012

Renan H. Cardoso received the B.Sc. degree in Computer Engineering from the Universidade Federal do Ceará (UFC), Brazil, in 2021. Since then, he has worked as a software engineer. In 2024, he completed his M.Sc. degree at UFC. His research interests include machine learning, tensor decompositions, and biomedical engineering.



C. Alexandre R. Fernandes received the B.Sc. degree in Electrical Engineering from the Universidade Federal do Ceara (UFC), Fortaleza, Brazil, in 2003, M.Sc. degrees from the UFC and University of Nice Sophia-Antipolis (UNSA), Nice, France, in 2005, and the double Ph.D. degree in signal processing from the UFC and UNSA, in 2009. He was a Teaching Assistant with the UNSA/FR (2008-2009) and Postdoctoral Fellow with the UFC (2009-2010). In 2010, he joined the UFC as a Full Professor with the Department of Computer Engineering. His research interests include machine learning, signal processing, tensor decompositions, biomedical engineering etc.

Pedro M. R. de Oliveira was born in Fortaleza, Brazil, in 1993. He received a B.Sc. degree in Computer Engineering from the Universidade Federal do Ceará (UFC), Brazil, in 2016, and an M.Sc. degree in Teleinformatics Engineering, also from UFC, in 2017. He received a Ph.D. degree in signal processing from the Université Côte d'Azur, France, in 2020, where he also worked as a Teaching Assistant. From Summer 2014 to Spring 2015, he was an Exchange Student at the Illinois Institute of Technology, U.S.A. During Summer 2015, he was an Intern Researcher at the George Washington University, U.S.A. From 2020, he works in the industry on the healthcare sector. His research interests include artificial inteligence, signal processing, tensor decompositions, applied mathematics, and biomedical engineering.