Computer Methods and Programs in Biomedicine 246 (2024) 108052

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.elsevier.com/locate/cmpb

Non-invasive estimation of atrial fibrillation driver position using long-short term memory neural networks and body surface potentials

Miriam Gutiérrez-Fernández-Calvillo [a,*], Miguel Ángel Cámara-Vázquez [a], Ismael Hernández-Romero [b], María S. Guillem [b], Andreu M. Climent [b], Carlos Fambuena-Santos [b], Óscar Barquero-Pérez [a]

[a] Rey Juan Carlos University, Cam. del Molino, 5, Madrid, 28942, Spain
[b] Universitat Politècnica de València, Camí de Vera s/n, València, 46022, Spain

ARTICLE INFO

Keywords:
Body surface potentials
Atrial fibrillation
Inverse problem
Recurrent neural network

ABSTRACT

Background and Objective: Atrial Fibrillation (AF) is a supraventricular tachyarrhythmia that can lead to thromboembolism, hearlt failure, ischemic stroke, and a decreased quality of life. Characterizing the locations where the mechanisms of AF are initialized and maintained is key to accomplishing an effective ablation of the targets, hence restoring sinus rhythm. Many methods have been investigated to locate such targets in a non-invasive way, such as Electrocardiographic Imaging, which enables an on-invasive and panoramic characterization of cardiac electrical activity using recording Body Surface Potentials (BSP) and a torso model of the patient. Nonetheless, this technique entails some major issues stemming from solving the inverse problem, which is known to be severely ill-posed. In this context, many machine learning and deep learning approaches aim to tackle the characterization and classification of AF targets to improve AF diagnosis and treatment.
Methods: In this work, we propose a method to locate AF drivers as a supervised classification problem. We employed a hybrid form of the convolutional-recurrent network which enables feature extraction and sequential data modeling utilizing labeled realistic computerized AF models. Thus, we used 16 AF electrograms, 1 atrium, and 10 torso geometries to compute the forward problem. Previously, the AF models were labeled by assigning each sample of the signals a region from the atria from 0 (no driver) to 7, according to the spatial location of the AF driver. The resulting 160 BSP signals, which resemble a 64-lead vest recording, are preprocessed and then introduced into the network following a 4-fold cross-validation in batches of 50 samples.
Results: The results show a mean accuracy of 74.75% among the 4 folds, with a better performance in detecting sinus rhythm, and drivers near the left superior pulmonary vein (R1), and right superior pulmonary vein (R3) whose mean sensitivity bounds around 84%-87%. Significantly good results are obtained in mean sensitivity (87%) and specificity (83%) in R1.
Conclusions: Good results in R1 are highly convenient since AF drivers are commonly found in this area: the left atrial appendage, as suggested in some previous studies. These promising results indicate that using CNN-LSTM networks could lead to new strategies exploiting temporal correlations to address this challenge effectively.

1. Introduction

Atrial Fibrillation (AF) is known as the most prevalent sustained cardiac rhythm disorder, which is associated with increased morbidity related to thromboembolism, heart failure, ischemic stroke and a decreased quality of life [22]. The incidence of AF is sharply increasing in developed countries due to the growing elderly population, affecting more than 33 million people in the world, and therefore entailing a major health problem [8].

AF treatments aim to restore sinus rhythm, either pharmacologically or by cardioversion. Nonetheless, the most effective approach includes surgery and catheter-based ablation, especially when managing symptomatic patients with paroxysmal or persistent AF intolerant to antiarrhythmic drug treatment [28].

* Corresponding author.
E-mail address: miriam.gutierrez@urjc.es (M. Gutiérrez-Fernández-Calvillo).

https://doi.org/10.1016/j.cmpb.2024.108052
Received 9 August 2023; Received in revised form 12 November 2023; Accepted 26 January 2024
Available online 1 February 2024

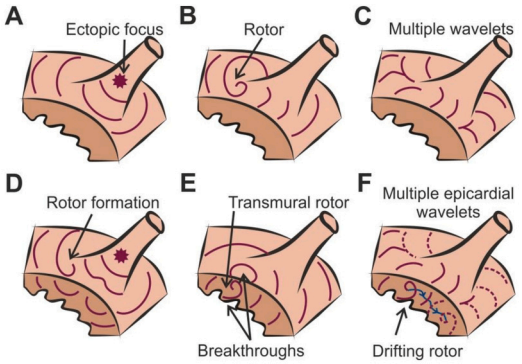

Fig. 1. Types of drivers mechanisms depend on complexity, waveform and propagation. AF drivers may follow one or a combination of the above electrical mechanisms when propagating along the atria tissue, namely ectopic focus, rotors (with rotational component) or multiple wavelets [30].

The mechanisms of both onset and maintenance of AF remain a matter of controversy. Several hypotheses have been proposed, namely atrial wavelets, macroreentries and localized sources (focal or reentrant, as shown in Fig. 1). Hence, to undertake an ablation procedure it is essential to characterize and locate these AF drivers (regions of the atria responsible for the AF), which can be accomplished through invasive electrophysiological (EP) studies or non-invasive methods [15]. During interventional invasive mapping, 8 to 128 simultaneous electrograms (EGM) are recorded simultaneously by introducing a catheter into the atrial chambers [25]. Nonetheless, the significant distance between catheter sensors (up to 1–2 cm) added to the complex atrial anatomy constrains the ability of invasive cardiac mapping to provide a precise spatial insight of AF substrate [18,27].

To address these limitations of traditional techniques, researchers have developed non-invasive techniques for mapping the atrial electrical activity, i.e. electrocardiographic imaging (ECGI), providing a higher spatial resolution. ECGI reconstructs the cardiac activity recording Body Surface Potentials (BSP) and a torso model of the patient [4]. Usually, BSPs are recorded using a vest with 250 electrodes, which in combination with geometrical information provided by a CT or MRI scan, allows reconstruction of the electrical potentials and repolarization patterns from the heart's surface [33]. However, the ECGI method faces the challenge of solving an ill-posed inverse problem due, in part, to the electrical attenuation between the epicardium and the surface of the torso [32]. Furthermore, the solution is unstable and can oscillate notably given small perturbations or noise in the geometrical models or the electrical input data [14]. As a consequence, we rely on the addition of constraints to reconstruct reliable and stable epicardial potentials from BSPs, being the most popular method Tikhonov regularization [35]. Accordingly, the use of ECGI raises some doubts on account of its mentioned limitations.

Machine Learning (ML) and Deep Learning (DL) have disrupted healthcare towards more powerful and effective diagnosis and treatment techniques, including AF characterization and detection. Deep Neural Networks (DNN), which are the simplest form of Neural Networks (NN), have been widely used to tackle many complex problems. [5] proposed a one-dimensional deep densely connected neural network classification, yielding promising results for AF diagnosis utilizing ten-second raw 12-lead ECG signals. Convolutional Neural Networks (CNN) are yet another popular architecture used for the first time for AF detection by [36], where Short-term Fourier transform (STFT) and stationary wavelet transform (SWT) were computed and fed to the network to solve a classification problem. Recurrent Neural Networks (RNN) were developed to address temporal correlations, like in [24], where ECG signals were fed to an attention mechanism of a bidirectional RNN (BiRNN). In the same line, Long-Short Term Memory Networks (LSTM) are also employed for AF detection in combination with CNN networks as in [1].

Several works address AF detection using ECG, yet few include the characterization and localization of AF drivers. [21] work used an encoder-decoder CNN structure to predict phase maps or driver core positions using images. Furthermore, [2] used autoencoders for heart surface potentials estimation from BSP and using CT images. A similar approach can be found in [23], where 12-lead ECG is employed to locate AF drivers using decision trees.

In this work, we improved previous approaches by proposing an hybrid architecture of CNN-LSTM which is able to exploit spatial and temporal correlations in the BSPs recordings to localize AF driver positions in the atria. [11,7,6].

The development and study of a non-invasive tool for mapping AF drivers is of great aid to electrophysiologists for electrical characterization, therapeutic guidance, and guidance in ablation therapy [16]. In this context, once the deep algorithms are correctly trained, predictions based on torso recordings are instantaneous. Precomputing a map of the electrical activity of the atria in real-time, fast and applicable to all patients without any risks is valuable to improve EP studies. These challenges have been attempted to be addressed by ECGI methods for years, yet remain insufficient. Nonetheless, deep learning methods have been demonstrated to have an exceptional capacity to solve complex problems, superior to the majority of conventional optimization and regularization techniques. More importantly, in this setup, the geometry of the patients is not necessary.

The next section will describe and detail computerized models, input data for training, and the designed architecture utilized for the experiments. Furthermore, performance metrics to measure the results will be displayed. Then, the results from the experiments will be shown. Finally, we extract some key conclusions and discuss the outcome of the work.

2. Methods

This section describes the materials and methods employed to undertake the experiments. First, computerized models are characterized, including the geometries of the torso and atria, the EGM signals and the atrial cell model description. Secondly, input signals, namely BSP signals, are defined. Then the implemented architecture is described along with the performance metrics utilized to measure the results and the experimental set-up is presented.

2.1. Computerized models

Taking into account a simplified endocardium-epicardium layer for the atrial tissue, a realistic 3D model of the atrial geometry was created using 284,578 nodes and 1,353,783 tetrahedrons (673.4 ± 0,001303 m between nodes) [10]. A total of 16 AF models were built from multiple complexity propagation patterns and driver locations. The AF driver can be found in the following anatomical regions: posterior left atrial wall (PLAW), left inferior pulmonary vein (LIPV), left superior pulmonary vein (LSPV), right inferior pulmonary vein (RIPV), right superior pulmonary vein (RSPV), right atrial appendage (RAA), and right atria free wall (RAFW). The resulting signals, which ranged from 2 to 5 seconds, were sampled at $fs = 500Hz$.

To achieve these conditions, four different types of propagation patterns were utilized to create the models, which are explained in the following paragraph. *Normal SR*, which resembles sinus rhythm, where atrial tissue is activated at 1.2 Hz. *RAA-located driver model*, with one functional re-entry at the RAA with atrial activation of 5 Hz at this singular point, while the rest of the tissue is activated at 3.1 Hz. *Two stable drivers (Two rotors, TR),* where two drivers are located in PLAW and SRAW regions. Furthermore, two anatomical re-entries are found in this model near to left PV and the right inferior pulmonary vein (RIPV). In this model, the activation ranges between 6-7 Hz. Finally, *the LSPV-located driver model*, where it can be found a single functional re-entry close to the LSTPV at a rotation rate of 7.8 Hz, while the rest of the

atrial tissue is activated at 3-6.5 Hz. In the last model, 50% of atrial cells are set under fibrotic conditions.

Final computerized models were generated from N = 2048 nodes for a single atria model, and M ranging from 2206 to 3970 nodes for each of the 10 torso models. By adding together all effective dipole contributions across the entire model, the models were developed under the premise of a homogeneous, unbounded, and quasi-static conducting medium.[12,11,29,31]. Hence, EGMs from the entire model were obtained as follows:

$$V(\vec{r}) = \sum_{\vec{r}} \frac{\vec{r}}{r^3} \cdot \vec{\nabla} V_m \tag{1}$$

where $V(\vec{r})$ is the EGM at the measuring point, V_m is the transmembrane potential distribution across the atria given $\vec{r}$, the distance between the measuring point and a point in the tissue domain, and r the corresponding scalar distance. Since transmembrane potentials were established in a dispersed 3D mesh, the gradient was produced by interpolating a quadratic function involving two nearby sites [20].:

$$V_{m,i} - V_{m,j} = c_1 x + c_2 y + c_3 z + c_4 x^2 + c_5 y^2 + c_6 z^2 + c_7 zy + c_8 yz + c_9 xz \tag{2}$$

where $V_{m,i}$ and $V_{m,j}$ are the transmembrane potentials at points i and j; x, y,z are the incremental Cartesian coordinates from j to i, and coefficients c_1 to c_9 were obtained through the least square method in, at least, nine neighbor points of each location.

Atrial cell model

Following the models proposed by [26] and [19], where the electrical activity of one myocyte is described in terms of its potentials and ionic currents as follows:

$$\frac{\partial V}{\partial t} = -\frac{I_{ion}}{C_m} \tag{3}$$

where V is the transmembrane potential, I_{ion} in the transmembrane ionic current and C_m corresponds to the cell membrane capacitance. Then, by incorporating the transmembrane currents produced by the intercellular Gap junction current (owing to the transmembrane potential gradient) in the earlier formulation, the electrical propagation across the atrial tissue was simulated:

$$\frac{\partial V_k}{\partial t} = -\frac{I_{ion}}{C_m} - \sum_{i=1}^{N} D_{K,i} \frac{V_k - V_i}{d_{k,i}^2} \tag{4}$$

where V_k is the transmembrane potential at node k, V_i is the transmembrane potential at the neighbor node i, $D_{k,i}$ corresponds to the diffusion coefficient between node k and i, and $d_{k,i}$ is the distance between the mentioned nodes. Given the anisotropic nature of atrial conduction, where the velocity increments in the longitudinal fiber orientation and decrease at transverse, the diffusion coefficient $D_{k,i}$ which modulates the intercellular ionic current follows the next expression:

$$D_{K,i} = D_{long} \cdot cos^2\alpha + D_{trans} \cdot sin^2\alpha \tag{5}$$

where α is the angle between longitudinal and transversal fibers, D_{long} and D_{trans} are the longitudinal and transverse diffusion coefficients, respectively for nodes k and i [30]. The modeling of fibrotic and scar tissue was achieved by setting the diffusion values of the involved nodes to zero. Regarding fibrotic tissue, a certain percentage of random nodes were settled to zero depending on the simulated propagation pattern. The final system of differential equations was solved by Runge-Kutta integration (using NVIDIA Tesla C2075 6G).

2.2. Input data

We propose to address the characterization of AF driver by locating its driver as a supervised classification problem. Hence, we divided the atria geometry into 7 regions, as depicted in Fig. 2, according to the odds of AF driver being allocated in each area [15]. Each region from 1-7 represents a class to which the AF driver belongs, assigning class 0 when no driver is found. Each time instant of the BSP was assigned one class from 0-7 following semi-automatic labeling which is visually inspected.

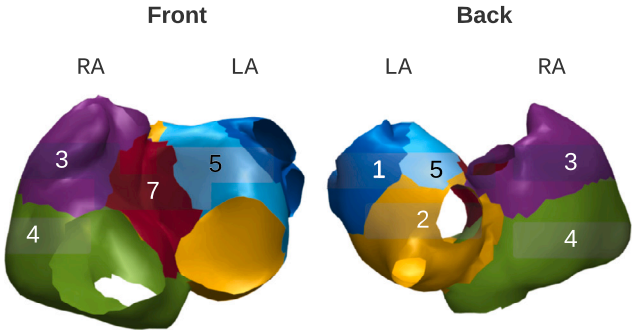

Fig. 2. Regions from 1 to 7 in which the location of the driver can be allocated are represented in the atria geometry. Each region is represented by a color and can be inspected by a front or back perspective of the atria. Regions 3 and 4 belong to the RA (Right Atria) whereas regions 1, 2, 5, and 7 lay within the LA (Left Atria) and septum area. Interestingly, no drivers were found in region 7 (septum) while labeling the data.

Simulated BSPs are computed from the 10 torso geometries and 16 AF models. To obtain the BSPs, the forward problem was solved using the A matrix of size MxN result of employing the boundary element method [3,9,29] as follows:

$$y_t = Ax_t \tag{6}$$

where y_t represents the BSPs for each time step. Reference to the Wilson Central Terminal (WCT) is then performed since real ECG recordings are referenced to this point on account of the electrical noise of the ground, and it can be modeled as:

$$ECG = ECG_{not\ ref} - ECG_{WCT} \tag{7}$$

where $ECG_{not\ ref}$ in the not referenced ECG, ECG_{WCT} i the WCT signal obtained as the average of the ECG at the WCT points. Thereby, applying the same reference to BSPs can be mathematically represented as:

$$y_{t,ref} = Ax_t - \frac{1}{N_{WCT}} \sum_{N \in WCT} Ax_t - \frac{1}{N_{WCT}} M_{WCT} x_t \tag{8}$$

thus

$$y_{t,ref} = \left(A - \frac{1}{N_{WCT}} M_{WCT} \right) x_t \tag{9}$$

where N_{WCT} is the number of WCT points, M_{WCT} is a null matrix except the rows corresponding to WCT leads, which have the same values as the equivalent rows of the A matrix [29]. Hence, to obtain the BSP referenced to the WCT, a corrected A_{WCT} matrix can be computed as:

$$A_{WCT} = A - \frac{1}{N_{WCT}} M_{WCT} \tag{10}$$

Subsequently, additive Gaussian noise is added to the obtained BSP with Signal-to-noise (SNR) of 20 dB. Then, signals are filtered using a fourth-order bandpass Butterworth filter (fc1 = 3 Hz and fc2 = 30 Hz) [11,29]. To obtain a more realistic representation of BPS recordings, 64 out of 192 nodes were selected from the computed BSPs. The selection was made to resemble the positioning of electrodes on a real chest vest, ensuring that the selected nodes are as close as possible to the real positions.

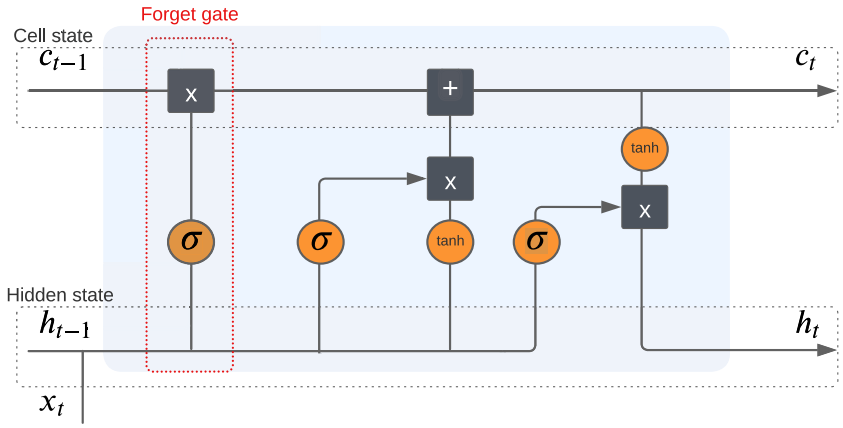

Fig. 3. LSTM cell unit. Information flows into the cell unit through the input gate, then the forget gate dictates which information from previous hidden states is passed to the next layers through the cell state [13].

2.3. RNN based NN architecture

RNN architecture

RNN are a family of NN for modeling sequence data [17]. Unlike multi-layer networks, RNN share parameters across the whole sequence model to accomplish this task. This approach enables a proper generalization across the different positions in time. In an RNN architecture, the information flows at time step t from the input layer $x^{<t>}$ to the output later $o^{<t>}$, and is stored at the hidden state $h^{<t>}$, which is the representation of the previous inputs. The hidden layer $h^{<t>}$ and the output $o^{<t>}$ are modeled as follows:

$$h^{<t>} = \sigma_1(U x^{<t>} + W h^{<t-1>} + b_1), \tag{11}$$

$$o^{<t>} = \sigma_2(V h^{<t>} + b_2) \tag{12}$$

where U, V and W represent the weights between the input and the hidden layers, the weights between two different time steps and the weights between the hidden and the output layer, respectively. On the other hand, b_1 and b_2 are the bias, meanwhile, σ_1 and σ_2 represent the activation functions. The output of a given time step is influenced by the previous hidden states as shown by equation (12).

However, during backpropagation exploding gradients or vanishing gradients issues may occur, causing the network to forget information located far in the past. This can happen on account of the uncontrolled growth or decrease in gradients as the backpropagation algorithm operates. To address this problem, LSTM networks were proposed to include mechanisms to selectively forget or update the information that is propagated across the network through the cell state, namely using forget gates [17].

Following the scheme shown in Fig. 3, first, the forget gate f_t decides which information flowing from the previous hidden state $h^{<t-1>}$ and based on the current input x_t is included in the cell state c_t as follows:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{13}$$

where σ is the sigmoid activation function, W_f are the weights associated with the forget gate, h_{t-1} is the previous hidden state, x_t is the current input and b_f is the bias.

Next, the input gate layer i_t decides which values from the cell state are updated, and a vector of new candidate values for the cell state $\tilde{C}_t$ using the tanh activation function is generated, following the next equations:

$$i_t = \sigma(W_i \cdot [h_t, x_t] + b_i), \tag{14}$$

$$\tilde{C}_t = tanh(W_C \cdot [h_t, x_t] + b_C) \tag{15}$$

To update the old cell state $C_{<t-1>}$ into the current cell state $C_{<t>}$, the following expression is followed:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t. \tag{16}$$

The output, which depends on the previous hidden state h_{t-1} and the current input x_t is obtained as follows:

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \tag{17}$$

Finally, the new current hidden state h_t is computed following the expression:

$$h_t = o_t \cdot tanh(C_t) \tag{18}$$

where the new hidden state is the pointwise multiplication of the updated cell state passed through a tanh activation function and the output.

CNN architecture

CNNs are a popular type of NN, mainly used in computer vision, which uses convolutional and pooling layers to extract features from input data. During convolutional layers, the convolution between different filters (kernels) and the input tensor is computed, extracting features from it and obtaining multiple feature maps.

First, the batch tensor is forwarded to the network through a convolutional layer. The convolution operation can be expressed as

$$a_{ij} = f\left(\sum_{m=0}^{s}\sum_{n=0}^{t} w_{m,n}\ x_{i+m,j+n} + b\right) \tag{19}$$

where a_{ij} is the activation on coordinates (i, j), $w_{m,n}$ is the coefficient of the filter on coordinates (m, n) $x_{i+m,j+n}$ is the value of the input tensor in coordinates $(i + m, j + n)$, b corresponds to the bias and $f(\cdot)$ represents a non-linear activation function.

The following layer, the pooling layer, reduces the spatial size of the convolutional layer output. These layers are often followed by a dense layer, hence the output of the convolutional network must be flattened first. There are many types of pooling layers, among the most popular ones stands max-pooling, which computes the largest value from each patch of the feature maps [13].

CNN-LSTM networks

In this work, we propose to use CNN-LSTM models, which are hybrid architectures that involve using CNN layers for extracting features from the input sequences combined with LSTM layers for signal classification. With this aim, the main sequence is split into blocks of smaller size which are fed to the 1-D convolutional networks, which will extract features from each of the subsequences. Following, LSTM layers will

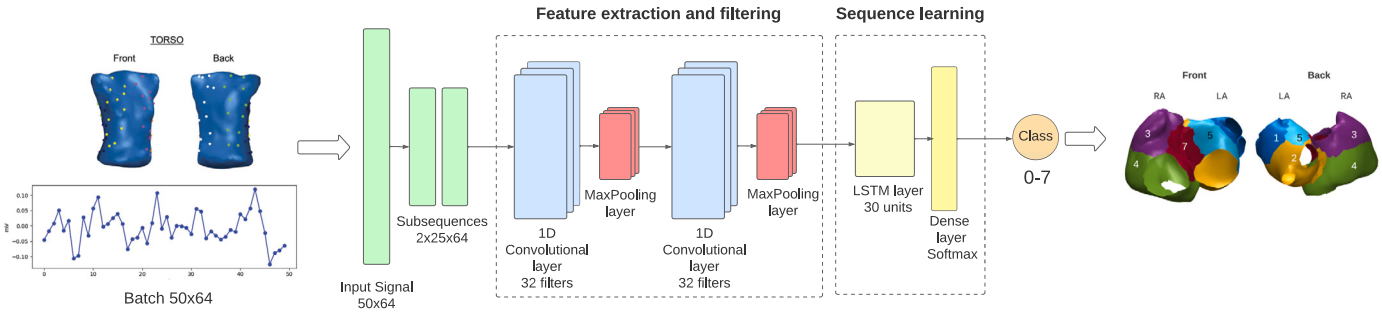

Fig. 4. CNN-LSTM Architecture. First the input BSP sets of size 50 x 64 (in this figure, only one of 64 BSPS is plotted for the sake of clarity) are divided into two subsequences of size 25 x 64. Then, the tensors are passed through a 1D Convolutional layer followed by pooling layers. Finally, the output of the CNN module is forwarded to an LSTM layer of 30 units and a softmax layer to perform the classification.

interpret the extracted features from each block and allow the sequence classification.

Aiming to fit this architecture layout, BSP signals are resampled from the original $fs = 500Hz$ to $fs = 50Hz$ before being introduced to the network; given that AF spectral activity is bounded between 2.5 and 25 Hz [34], feeding the network with signals sampled to 500 Hz would result in redundancy between sequential batches. To facilitate adequate training and feature extraction from the sequences, we resample BSP to lower frequencies aiming to target AF phenomenon.

Then, the input BSPM segments of size 50 time steps by 64 features were split into two blocks of size 25 time steps. As Fig. 4 shows, the subsequences are passed through two 1-dimensional convolutional layers of 32 filters as well as max pooling layers of size 5 and 3. The output is then flattened and passed to a 30-unit LSTM layer. The last layer is a fully connected layer with a softmax activation function which allows to perform an 8-class classification. Multiple dropouts layers are introduced with the objective of avoiding overfitting with a probability between 0.5 and 0.8. Moreover, L2 regularization is applied in the last layer.

2.4. Performance metrics

Five different metrics are employed during the assessment of the models:

- Accuracy (Acc). Used as a global metric, it yields the proportion of correctly predicted driver regions to the total observations:

$$Acc = \frac{TP + TN}{Total} \tag{20}$$

  where TP (true positives) are the count of drivers that are correctly classified in a given region whereas TN (true negatives) refers to the number of drivers that do not belong to a certain region and the machine predicts so.
- Sensitivity (or True positive rate, TPR). It represents the percentage of detected drivers for a given region that are correctly classified:

$$Sensitivity = \frac{TP}{TP + FN} \tag{21}$$

  where TP (true positives) are the count of drivers that are correctly classified in a given region whereas FN (false negatives) are the drivers that are incorrectly classified into another region. A high sensitivity signifies a good performance of the test identifying true positives.
- Specificity (or True negative rate, TNR). It is the percentage of not detected drivers for a given region that are correctly classified.

$$Specificity = \frac{TN}{TN + FP} \tag{22}$$

  where TN (true negatives) refers to the number of drivers that do not belong to a certain region and the machine predicts so, meanwhile, FP (false positives) is the number of drivers classified in a given region, but actually belonging to another class. A high specificity denotes a good performance of the test identifying true negatives.
- Precision. It provides insight into all drivers classified in a region, and how many are actually located in that region:

$$Precision = \frac{TP}{TP + FP} \tag{23}$$

  where TP (true positives) is the count of drivers that are correctly classified in a given region; FP (false positive) is the number of drivers classified in a given region, but actually belonging to another class; TN (true negatives) refers to the number of drivers that do not belong to a certain region and the machine predicts so; FN (false negatives) are the drivers that are incorrectly classified into another region.

2.5. Experimental set-up

The resulting dataset is comprised of 160 BSPs sets (derived from 12 AF models and one sinus rhythm model) concatenated sequentially constituting a matrix of size (29,000 time instants × 64 nodes). Each time step of the BSP signals is associated with a given position of the driver in the atria (from 0 to 7). Given the amount of data available, regions 2 and 3 were merged on account of their anatomical proximity, aiming to improve label balance, as shown in Fig. 6.

With the objective of performing a fair validation, a 4-fold cross-validation scheme was chosen. Each BSP signal vector was split into train and test batches of size 50 samples, analogous to one second of the sequence (obtaining blocks of size 50 x 64), considering consecutive time steps aiming to preserve temporal correlation. Then, this process was repeated 4 times with different test subsets with no group overlapping. The networks are trained following a many-to-one scheme, denoting that each input batch of 50 samples is assigned only one label as an output (the mode of the 50 samples). The final performance of the models is obtained by computing the mean and standard deviation of the 4 folds as depicted in Fig. 5.

3. Results

In this section, the performance metrics evaluated in the test subset for each region of the classification are presented.

The mean and standard deviation of the following metrics for the 4 folds are computed: accuracy, precision, sensitivity, and specificity. The driver is correctly located on average for the 4 fold 74.75% ± 6.02 of the times, suggesting a low variation among folds.

However, considering overall results may not be suitable for this specific problem, since performance for each region may vary considerably. Hence each region's metrics are computed and reported in Fig. 7, showing that the detection of drivers in R1 is, together with sinus rhythm (no

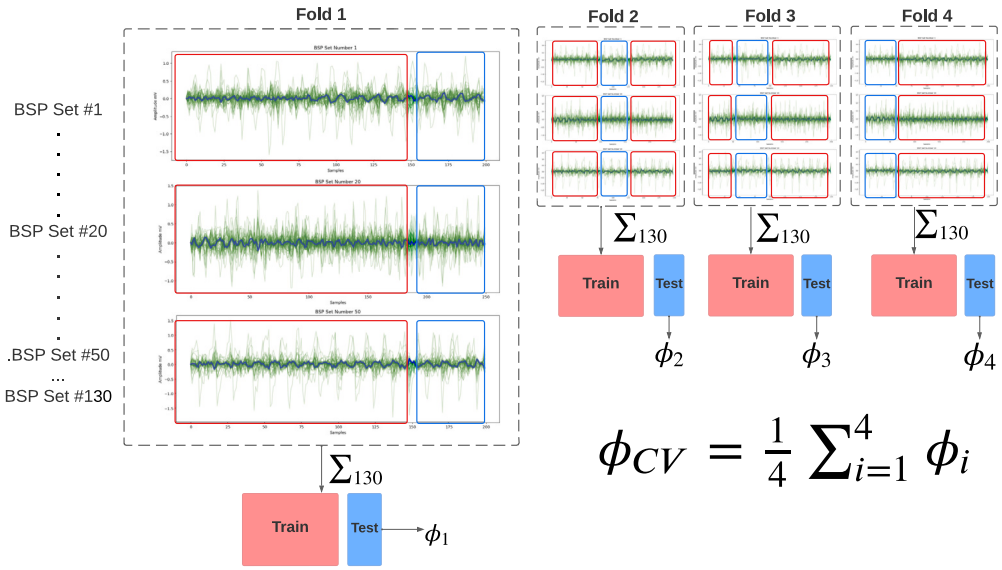

Fig. 5. Scheme followed during the CV. In each fold, a different segment of each of the 160 BSPS is selected as the test subset. Then, all the segments stemming from the 160 sets are merged to conform to a final train and test subsets. Algorithms are trained using train and metrics named as ϕ are obtained from the test subset. Finally, the mean and standard deviation of the 4 folds is computed. The figure depicts an example for 3 different BSPS sets: number 1, 20 and 50.

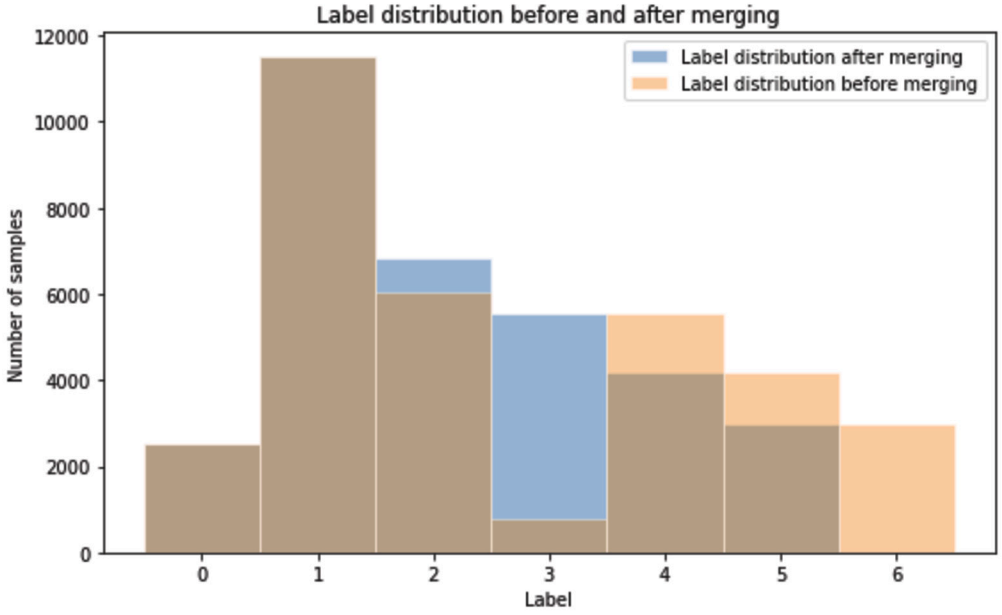

Fig. 6. Histogram of labels before merging and after merging regions 2 and 3. Underrepresented region 3 is merged with region 2, which is anatomically close to the former.

driver), the class with the highest number of existing drivers detected correctly classified (true positives). However, R1 drivers are occasionally mistaken for R2, which can be explained by the proximity of these regions in the atria geometry. On the other hand, the regions that show the lowest performance are R2 and R4, exhibiting the highest dispersion in the number of existing drivers that are detected and correctly classified (true positives).

Fig. 8 depicts the performance in terms of sensitivity (percentage of drivers located in a region being correctly located in the given class), specificity (percentage of drivers not located in a region being correctly classified as such) classified in each region and precision (measuring how many drivers are actually located in a given region from all drivers classified in that region). The former metric is the highest in class no driver and R1 with a mean percentage of 84% and 87% respectively, with little dispersion, unlike R2, R4 and R5 which report a wider interquartile range. Addressing the specificity, better results are obtained with class no driver, R2, R4 and R5. In this case, the median of the specificity for R3 and R4 (which have a worse overall performance) is slightly higher than in R1, this may be due to imbalances found in the dataset. The percentage of times the machine was correct when predicting a driver in its true region (precision) peaks in class R0 and R1, with little dispersion and medians above 80%, unlike R2, which shows the lowest mean value (47%).

According to the curves shown in Fig. 9, the overall estimated percentage of correctly classified drivers bounds between 71% and 82%.

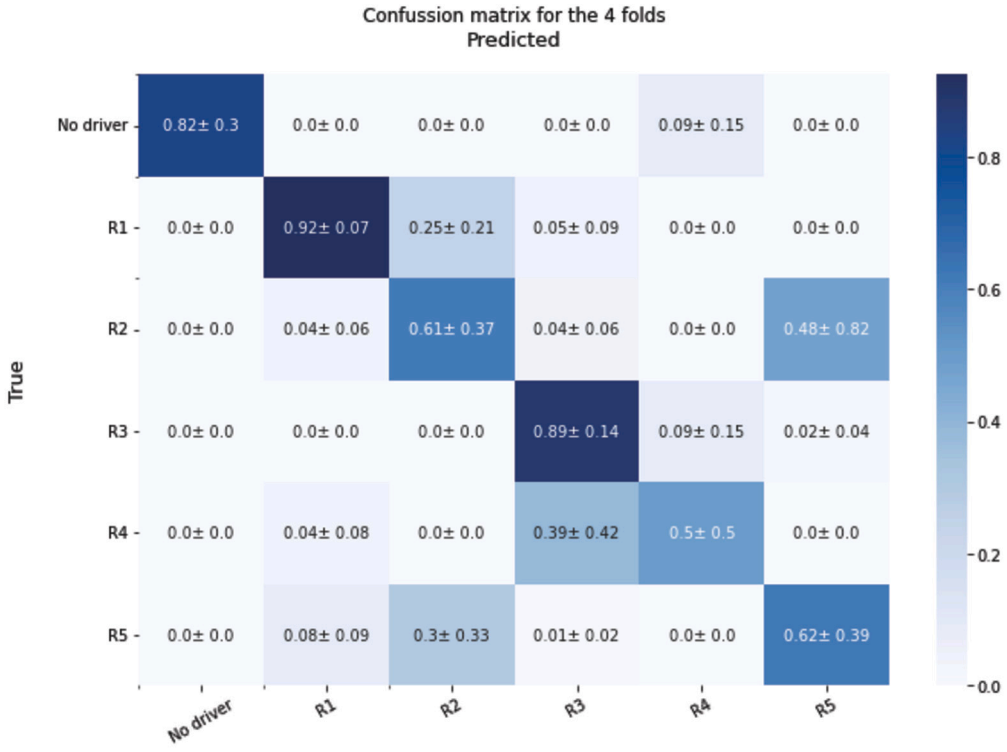

Fig. 7. Confusion Matrix for the 4 folds: the mean and standard deviation of each class is averaged for the 4 folds.

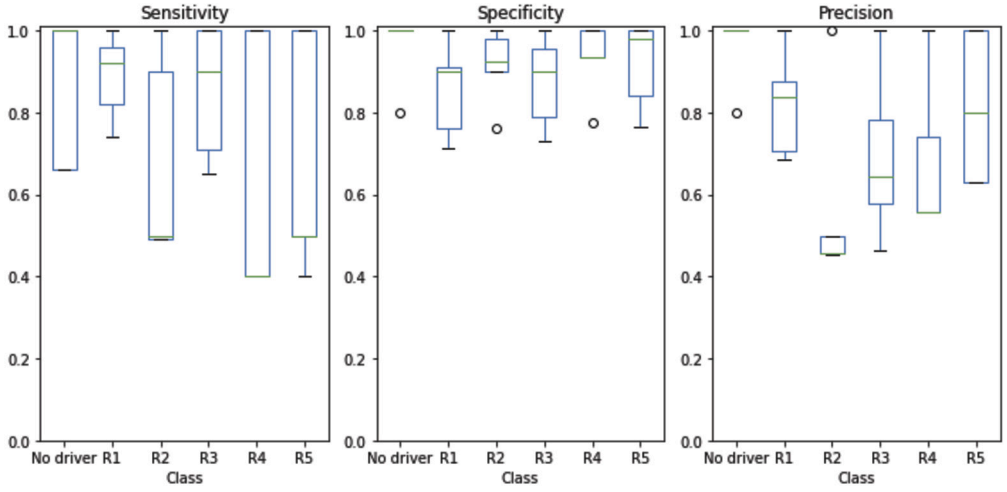

Fig. 8. Box plots representing sensitivity, specificity and precision in each region. The best performance is obtained with R0 and R1.

The train curve progresses parallel to the test curve and does not show drastic signs of overfitting.

4. Discussion

In this section, we discuss the obtained results and some limitations encountered during the process are presented along with some proposals to improve and optimize the experiments and scope.

The proposed CNN-LSTM architecture methodology outperforms previous works [6] for AF driver detection based on regular CNN architectures, demonstrating a positive impact of the inclusion of LSTM layers to analyze temporal correlations in the signals. This method has provided very promising results for 6 regions of the atria and 20 dB noise level in terms of precision, where the mean precision among regions is above 70% in 4 of the 6 regions, providing reliable predictions for most of the regions. The good results in sensitivity, with a mean value bounded between 67% and 90%, show a major ability of the tool to predict correctly the location of the driver if it occurs in the majority of atrial regions, which accounts for a valuable predicting feature in the clinical environment.

The model is able to almost perfectly detect sinus rhythm, and very good results are obtained in R1 as well, which is very convenient and highly remarkable since persistent AF drivers are commonly found in this area: the left atrial appendage, as previous studies reveal [15]. On the contrary, detection in R5 and R4 are inferior, caused by a minor incidence of occurrence of AF drivers in these regions.

Some limitations encountered during the development of the experiments are directly related to data scarcity and data imbalance. Due to the limited dataset size, acquiring generalization and robustness is a challenging milestone. Several technical resources have been deployed aiming to solve this problem, although the robustness and performance may be highly benefited from the inclusion of more train models in the

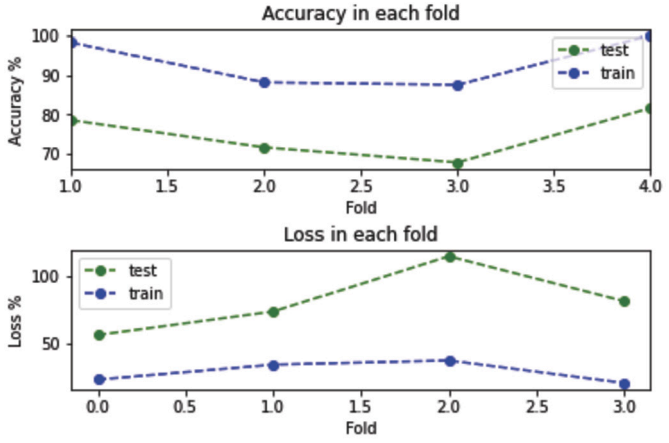

Fig. 9. Accuracy and loss are computed from the train and test subsets for each fold. The accuracy results for the test set in the different folds range from 71% to 82%.

dataset. Moreover, class imbalance issues have been also experienced as well. This imbalance is inherent in this problem: the natural electrophysiological propagation and anatomical conditions have promoted that the finding of these rotors is naturally more prominent in certain regions of the atria. Thus, we consider that the intrinsic imbalance of the classes is a depiction of the actual electrical propagation patterns, and, although it is not beneficial for the training of the algorithm, may not be fair to radically correct the phenomena.

From this point, several pathways can be led to improve the results. Firstly, including more AF models and geometries is essential to ensure adequate training and generalization. Ideally, in the future, real data from patients could be employed to train the algorithms. Nonetheless, the lack of gold-standard or existing labeled databases makes this task challenging.

Moreover, the inclusion of new AF models would enable the implementation of more complex architectures, such as autoencoders, which would be useful to extract important insights from the BSPS signals in order to predict their behavior.

5. Conclusions

In this section, the conclusions and impact of the results are presented.

Developing a robust, non-invasive mapping technique for intracardiac activity could greatly enhance the efficiency of electrophysiologists' work, enabling safer, quicker and more precise interventions in patients with AF. The utilization of deep learning methodologies presents an interesting approach to tackle this challenge, as they have proven their capability to handle intricate tasks within analogous clinical scenarios. Notably, their adaptability is a key asset; within the same architecture, these methodologies can be refined as the dataset expands. This is true as long as the quality of data is preserved in the new examples, and the new cases are representative and relevant for the task.

The presented results confirm that deep architectures are feasible methodologies for addressing this problem, more specifically, the use of the LSTM layer seems to help the models to capture AF activity, which appears to be important for locating the AF driver. This is reflected by the fact that the results outperform previous works based only on traditional convolutional and fully connected layers. The model can, in a way, detect if there is fibrillation in addition to locating the region where the source of the AF is, which is revealing for the design and development of future deep learning-enhanced tools.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work has been partially supported by: Ministerio de Ciencia e Innovación (PID2019-105032GB-I00, PID2022-136887NB-I00), Rey Juan Carlos University (2023/00004/006-F918), Instituto de Salud Carlos III and Ministerio de Ciencia, Innovación y Universidades (supported by FEDER Fondo Europeo de Desarrollo Regional PI17/01106 and Consejería de Ciencia, Universidades e Innovación of the Comunidad de Madrid through the program RIS3 (S-2020/L2-622), EIT Health Activity code 19600 and the European Union Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 860974.

References

[1] R.S. Andersen, A. Peimankar, S. Puthusserypady, A deep learning approach for real-time detection of atrial fibrillation, Expert Syst. Appl. 115 (2019) 465–473.
[2] T. Bacoyannis, J. Krebs, N. Cedilnik, H. Cochet, M. Sermesant, Deep learning formulation of ecgi for data-driven integration of spatiotemporal correlations and imaging information, in: International Conference on Functional Imaging and Modeling of the Heart, Springer, 2019, pp. 20–28.
[3] R.C. Barr, M. Ramsey, M.S. Spach, Relating epicardial to body surface potential distributions by means of transfer coefficients based on geometry measurements, IEEE Trans. Biomed. Eng. (1977) 1–11.
[4] L.R. Bear, O. Bouhamama, M. Cluitmans, J. Duchateau, R.D. Walton, E. Abell, C. Belterman, M. Haissaguerre, O. Bernus, R. Coronel, et al., Advantages and pitfalls of noninvasive electrocardiographic imaging, J. Electrocardiol. 57 (2019) S15–S20.
[5] W. Cai, Y. Chen, J. Guo, B. Han, Y. Shi, L. Ji, J. Wang, G. Zhang, J. Luo, Accurate detection of atrial fibrillation from 12-lead ecg using deep neural network, Comput. Biol. Med. 116 (2020) 103378.
[6] M.Á. Cámara-Vázquez, I. Hernández-Romero, E. Morgado-Reyes, M.S. Guillem, A.M. Climent, O. Barquero-Pérez, Non-invasive estimation of atrial fibrillation driver position with convolutional neural networks and body surface potentials, Front. Physiol. 12 (2021).
[7] M.Á. Cámara-Vázquez, A. Oter-Astillero, I. Hernández-Romero, M. Rodrigo, E. Morgado-Reyes, M.S. Guillem, A.M. Climent, Ó. Barquero-Pérez, Atrial fibrillation driver localization from body surface potentials using deep learning, in: 2020 Computing in Cardiology, IEEE, 2020, pp. 1–4.
[8] S.S. Chugh, R. Havmoeller, K. Narayanan, D. Singh, M. Rienstra, E.J. Benjamin, R.F. Gillum, Y.H. Kim, J.H. McAnulty Jr, Z.J. Zheng, et al., Worldwide epidemiology of atrial fibrillation: a global burden of disease 2010 study, Circulation 129 (2014) 837–847.
[9] J. De Munck, A linear discretization of the volume conductor boundary integral equation using analytically integrated elements (electrophysiology application), IEEE Trans. Biomed. Eng. 39 (1992) 986–990.
[10] O. Dössel, M.W. Krueger, F.M. Weber, M. Wilhelms, G. Seemann, Computational modeling of the human atrial anatomy and electrophysiology, Med. Biol. Eng. Comput. 50 (2012) 773–799.
[11] C. Figuera, V. Suárez-Gutiérrez, I. Hernández-Romero, M. Rodrigo, A. Liberos, F. Atienza, M.S. Guillem, Ó. Barquero-Pérez, A.M. Climent, F. Alonso-Atienza, Regularization techniques for ecg imaging during atrial fibrillation: a computational study, Front. Physiol. 7 (2016) 466.
[12] V.M. García-Mollá, A. Liberos, A. Vidal, M. Guillem, J. Millet, A. Gonzalez, F.J. Martínez-Zaldívar, A.M. Climent, Adaptive step ode algorithms for the 3d simulation of electric heart activity with graphics processing units, Comput. Biol. Med. 44 (2014) 15–26.
[13] I. Goodfellow, Y. Bengio, A. Courville, Deep Learning, MIT Press, 2016.
[14] R. Gulrajani, The forward and inverse problems of electrocardiography, IEEE Eng. Med. Biol. Mag. 17 (1998) 84–101, https://doi.org/10.1109/51.715491.
[15] M. Haissaguerre, M. Hocini, A. Denis, A.J. Shah, Y. Komatsu, S. Yamashita, M. Daly, S. Amraoui, S. Zellerhoff, M.Q. Picat, et al., Driver domains in persistent atrial fibrillation, Circulation 130 (2014) 530–538.
[16] I. Hernandez-Romero, R. Molero, C. Fambuena-Santos, C. Herrero-Martin, A.M. Climent, M.S. Guillem, Electrocardiographic imaging in the atria, Med. Biol. Eng. Comput. 61 (2023) 879–896.
[17] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Comput. 9 (1997) 1735–1780.
[18] Z.F. Issa, J.M. Miller, D.P. Zipes, Chapter 6 - Advanced mapping and navigation modalities, in: Z.F. Issa, J.M. Miller, D.P. Zipes (Eds.), Clinical Arrhythmology and Electrophysiology: A Companion to Braunwald's Heart Disease, second edition, W.B. Saunders, Philadelphia, 2012, pp. 111–143.

[19] J.T. Koivumäki, G. Seemann, M.M. Maleckar, P. Tavi, In silico screening of the key cellular remodeling targets in chronic atrial fibrillation, PLoS Comput. Biol. 10 (2014) e1003620.
[20] C.L. Lawson, C^1 surface interpolation for scattered data on a sphere, Rocky Mt. J. Math. (1984) 177–202.
[21] J. Lebert, N. Ravi, F.H. Fenton, J. Christoph, Rotor localization and phase mapping of cardiac excitation waves using deep neural networks, Front. Physiol. 12 (2021).
[22] G.Y. Lip, H.F. Tse, Management of atrial fibrillation, Lancet 370 (2007) 604–618.
[23] G. Luongo, L. Azzolin, M.W. Rivolta, R. Sassi, J.P. Martínez, P. Laguna, O. Dössel, A. Loewe, Non-invasive identification of atrial fibrillation driver location using the 12-lead ecg: pulmonary vein rotors vs. other locations, in: 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), IEEE, 2020, pp. 410–413.
[24] S. Mousavi, F. Afghah, U.R. Acharya, Han-ecg: an interpretable atrial fibrillation detection model using hierarchical attention networks, Comput. Biol. Med. 127 (2020) 104057.
[25] S.M. Narayan, D.E. Krummen, K. Shivkumar, P. Clopton, W.J. Rappel, J.M. Miller, Treatment of atrial fibrillation by the ablation of localized sources: confirm (conventional ablation for atrial fibrillation with or without focal impulse and rotor modulation) trial, J. Am. Coll. Cardiol. 60 (2012) 628–636.
[26] A. Nygren, C. Fiset, L. Firek, J.W. Clark, D.S. Lindblad, R.B. Clark, W.R. Giles, Mathematical model of an adult human atrial cell: the role of k+ currents in repolarization, Circ. Res. 82 (1998) 63–81.
[27] T. Oesterlein, D. Frisch, A. Loewe, G. Seemann, C. Schmitt, O. Dössel, A. Luik, Basket-Type Catheters: Diagnostic Pitfalls Caused by Deformation and Limited Coverage, BioMed Research International, 2016.
[28] R. Parameswaran, A.M. Al-Kaisey, J.M. Kalman, Catheter ablation for atrial fibrillation: current indications and evolving technologies, Nat. Rev. Cardiol. 18 (2021) 210–225.
[29] J. Pedrón-Torrecilla, M. Rodrigo, A.M. Climent, A. Liberos, E. PéREZ-DAVID, J. Bermejo, A. Arenal, J. Millet, F. Fernández-Avilés, O. Berenfeld, et al., Noninvasive estimation of epicardial dominant high-frequency regions during atrial fibrillation, J. Cardiovasc. Electrophysiol. 27 (2016) 435–442.
[30] M. Rodrigo, Non-invasive identification of atrial fibrillation drivers, Ph.D. thesis. Tesis doctoral, Universitat Politècnica de València, Valencia, España, 2016.
[31] M. Rodrigo, A.M. Climent, A. Liberos, F. Fernández-Avilés, O. Berenfeld, F. Atienza, M.S. Guillem, Technical considerations on phase mapping for identification of atrial reentrant activity in direct-and inverse-computed electrograms, Circ. Arrhythm. Electrophysiol. 10 (2017) e005008.
[32] M. Rodrigo, M.S. Guillem, A.M. Climent, J. Pedrón-Torrecilla, A. Liberos, J. Millet, F. Fernández-Avilés, F. Atienza, O. Berenfeld, Body surface localization of left and right atrial high-frequency rotors in atrial fibrillation patients: a clinical-computational study, Heart Rhythm 11 (2014) 1584–1591, https://doi.org/10.1016/j.hrthm.2014.05.013, focus Issue: Atrial Fibrillation.
[33] Y. Rudy, Noninvasive ecg imaging (ecgi): mapping the arrhythmic substrate of the human heart, Int. J. Cardiol. 237 (2017) 13–14.
[34] M. Stridh, A. Bollmann, S.B. Olsson, S. Leif, et al., Detection and feature extraction of atrial tachyarrhythmias, IEEE Eng. Med. Biol. Mag. 25 (2006) 31–39.
[35] A.N. Tikhonov, V.J. Arsenin, V.I. Arsenin, V.Y. Arsenin, et al., Solutions of Ill-Posed Problems, Vh Winston, 1977.
[36] Y. Xia, N. Wulan, K. Wang, H. Zhang, Detecting atrial fibrillation by deep convolutional neural networks, Comput. Biol. Med. 93 (2018) 84–92.