DE GRUYTER

Biomed. Eng.-Biomed. Tech. 2015; 60(3): 245–255

Christopher Schilling, Matthias Keller, Daniel Scherr, Tobias Oesterlein,
Michel Haïssaguerre, Claus Schmitt, Olaf Dössel and Armin Luik*

Fuzzy decision tree to classify complex fractionated atrial electrograms

Abstract: Catheter ablation has emerged as an effective treatment strategy for atrial fibrillation (AF) in recent years. During AF, complex fractionated atrial electrograms (CFAE) can be recorded and are known to be a potential target for ablation. Automatic algorithms have been developed to simplify CFAE detection, but they are often based on a single descriptor or a set of descriptors in combination with sharp decision classifiers. However, these methods do not reflect the progressive transition between CFAE classes. The aim of this study was to develop an automatic classification algorithm, which combines the information of a complete set of descriptors and allows for progressive and transparent decisions. We designed a method to automatically analyze CFAE based on a set of descriptors representing various aspects, such as shape, amplitude and temporal characteristics. A fuzzy decision tree (FDT) was trained and evaluated on 429 predefined electrograms. CFAE were classified into four subgroups with a correct rate of 81±3%. Electrograms with continuous activity were detected with a correct rate of 100%. In addition, a percentage of certainty is given for each electrogram to enable a comprehensive and transparent decision. The proposed FDT is able to classify CFAE with respect to their progressive transition and may allow objective and reproducible CFAE interpretation for clinical use.

Keywords: ablation; atrial electrograms; atrial fibrillation; CFAE; classification; fuzzy decision tree.

*Corresponding author: Armin Luik, Städtisches Klinikum
Karlsruhe, Karlsruhe, Germany,
E-mail: Armin.Luik@klinikum-karlsruhe.de
**Christopher Schilling, Matthias Keller, Tobias Oesterlein and
Olaf Dössel:** Institute of Biomedical Engineering, Karlsruhe Institute
of Technology (KIT), Karlsruhe, Germany
Daniel Scherr and Michel Haïssaguerre : Hôpital Cardiologique du
Haut-Lévêque, Bordeaux-Pessac, France
Claus Schmitt: Städtisches Klinikum Karlsruhe, Karlsruhe, Germany

Introduction

Atrial fibrillation (AF) is the most common cardiac arrhythmia, which affects morbidity and mortality [6]. AF is therefore a major and growing expense for health systems [11]. Catheter ablation has emerged as an effective treatment strategy in the last years. Since the discovery of foci inside the pulmonary veins (PVs) as a trigger for paroxysmal AF in 1998 [12], the technique of pulmonary vein isolation (PVI) has been established with acceptable success rates [29]. However, clinical outcome of PVI in persistent and long standing persistent AF is poor. This suggests different underlying mechanisms [4, 39, 40] outside the PV regions: multiple random propagating wavelets, focal electrical discharges or breakthroughs, and localized re-entrant activity with fibrillatory conduction [5]. However, differentiation during an electrophysiological (EP) study is limited as specific electrogram characteristics are not known. Ndrepepa et al. showed that persistent AF presents with shorter cycle lengths and more disorganized activity than paroxysmal AF [27]. Areas with complex fractionated atrial electrograms (CFAE) have been reported to potentially represent AF substrate sites [1, 24]. Nademanee defined CFAE as fractionated electrograms composed of two deflections or more, and/or as the perturbation of the baseline with continuous deflection of a prolonged activation complex and a median cycle length of <120 ms over a 10 s recording period [24]. This author was the first to address CFAE as a target for catheter ablation with respectable success rates [24], but their results have not been reproduced by other groups [32, 33]. At present, the definitions used for CFAE have been variable and our understanding of its mechanistic significance remains incomplete [20]. It has been shown that both, the prevalence and distribution of CFAE differed significantly when different CFAE definitions were utilized [21, 36, 43]. Furthermore, filter settings in clinical signal acquisition systems can strongly influence the appearance of measured signals [14]. Although, the integration of automatic CFAE detection algorithms in 3D mapping systems has facilitated CFAE site retrieval, they vary in definition and do not respect the progressive transition between the

CFAE classes [33]. It is therefore of importance to work out the different variants of CFAE and to develop mathematical algorithms which respect the fuzziness of these electrograms.

Wells et al. classified bipolar atrial electrograms during AF into four groups. Whereas type 1 and type 2 describes more regularized AF, type 3 is completely disorganized and type 4 is anything in between [42]. Allessie et al. described the complete activation pattern of the atrial wall obtained by high resolution mapping, and classified AF types according to the number of wavelets activating the atrial wall [1].

The detection of signals containing CFAE patterns has been addressed in literature. Cuesta-Frau et al. [9] used sample entropy to differentiate between CFAE and non-CFAE signals despite the presence of measuring artifacts. Recurrence quantification analysis was successfully applied by Navoret and colleagues [26] in the same context. Recurring patterns in CFAE were focused upon by Ciaccio et al. [8], resulting in an approach using transform coefficients on synchronized AF data.

There are some atrial electrogram decision making algorithms described in the literature [2, 18, 30] which are mainly based on the Wells classification. Kordik [18] defined an array of signal features which leads to a measure of the fractionation of a given signal. A neural net classifies the signals into four CFAE classes. They worked on a database with 113 annotated electrograms. Nollo et al. [30] defined features to describe the state of organization of CFAE. The best features were chosen on the basis of their Jeffries-Matusita distance. In a final step, a support-vector-machine classifies CFAE into three classes. Their database included 100 annotated electrograms. The limitations of these algorithms are, that they do not deliver a measure of certainty for the chosen classification.

The aim of this study was to develop an algorithm which addresses the fuzziness, i.e., the progressive transition between the CFAE classes, and presents a percentage of certainty of the selected subgroups. This will objectivize CFAE interpretation in a more reproducible and transparent way.

Methods

Study population

The study cohort consisted of 11 patients who underwent catheter ablation of AF. The local Ethics Committee approved this study (according to the declaration of Helsinki), and all patients gave written informed consent. Eight patients had persistent and three

paroxysmal AF. During the EP study, the left atrium was reconstructed using a 3D mapping system (Ensite NavX, St. Jude Medical, St. Paul, MN, USA) and PVI was performed. After PVI, bipolar electrograms were recorded using circular multipolar mapping catheters (10 polar Lasso, Biosense Webster, Diamond Bar, CA, USA; 14 polar OrbiterPV, Bard Electrophysiology, Lowell, MA, USA; 14 polar Optima, St. Jude Medical, St. Paul, MN, USA). After the procedure the electrograms and xyz-coordinates were exported and retrospectively analyzed. All together 605 recordings were analyzed. Each had a length of 5 s and was sampled at 1.2 kHz. The data was prefiltered by the measurement system with a high pass at 30 Hz and a low pass at 300 Hz.

For automatic classification, the Wells' criteria [42] were modified by changing type 3 and 4. This led to a classification with a continuous increase in complexity (Figure 1).

Data preprocessing

All electrograms were preprocessed in the same way. Baseline wandering and low-frequency noise were extracted from the atrial electrograms using a discrete wavelet transform-based approach [16]. Based on the sample frequency (f_s) the signal is decomposed up to level $n = \mathrm{ld}(f_s)$ and the approximation of level n is set to zero. After this step, the signal is composed again. The high frequency disturbance is removed by a conventional Butterworth low pass filter of order 4 with a cut off frequency of $f_{LP} = 300$ Hz [15]. After removing the noise by means of signal processing, the 605 stationary 5 s signals were classified by two physicians from different centers according to the above predefined classes (see Table 1). After the classification process, only electrograms with an unambiguous assignment were used to train and test the fuzzy decision tree (FDT). So finally the database consisted of 429 classified electrograms.

Descriptors of CFAE

To optimize the automatic electrogram characterization multiple features were generated. In the following sections a list of descriptors is delineated, reflecting the mathematical characteristics of CFAE.

Time domain descriptors based on non-linear energy operator: In sinus rhythm or other regularized atrial activity, intracardiac recordings are only presenting deflections if an excitation wave front propagates near the recording electrode. At baseline, no electrical activity is present. Therefore, the electrograms can be divided into segments with and without baseline crossings. Segments with baseline crossings were defined as active. Localization of active segments during AF is complicated, because they vary in length and number. A mathematical expression of active segments can be described as the energy, which forms the basis for time domain analysis.

Teager's non-linear energy operator [NLEO, (1)] [38] forms the basis of this descriptor. Adding an adaptive threshold on the low pass filtered NLEO enabled the separation of active and inactive segments [28, 38]. The output of the NLEO E_n can be considered as an indication of the energy of the signal $x(n)$ and it is proportional to the frequency and amplitude of the signal. Based on this segmentation a set of descriptors is defined.
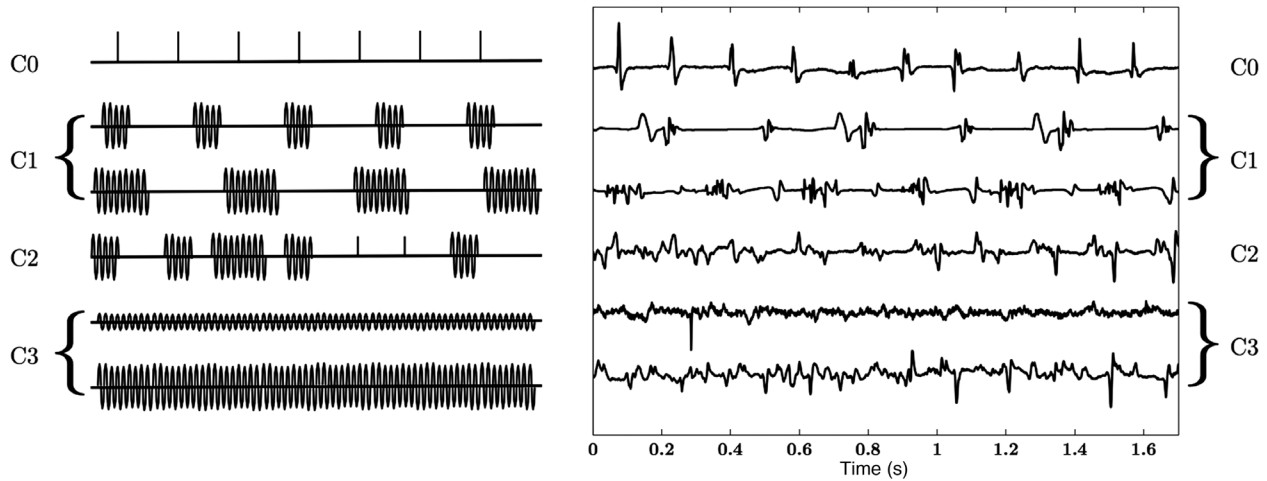
$$E_n = x_n^2 - x_{n+1} x_{n-1} \tag{1}$$

Figure 1: CFAE during atrial fibrillation.
Left: schematic presentation of the four variants according to the modified classification by Wells et al. Right: recorded electrograms during AF. C0 is a non-fractionated atrial electrogram of high frequency, C1 is a fractionated atrial electrogram with periodic activity, C2 is a mixture of periodic fractionated and periodic non-fractionated atrial electrograms, and C3 is a high frequency atrial electrogram with continuous activity.

The sum of the length of all active segments during a 5 s recording divided by the total length of the signal is called "activity ratio" (AR, descriptor D10).

$$AR = \frac{1}{L}\sum_{i=1}^{M} l_i \qquad (2)$$

with L the length of the total signal, M the number of active segments and l_i the length of an active segment i. Other descriptors are the mean length of active segments (MLAS, descriptor D11) and the standard deviation of active segments (sdMLAS, descriptor D12). They are defined as

$$MLAS = \frac{1}{M}\sum_{i=1}^{M} l_i \quad sdMLAS = \left(\frac{1}{M}\sum_{i=1}^{M}(l_i - MLAS)^2\right)^{1/2} \qquad (3)$$

Five more descriptors are defined based on the found active segments. These are the number of active segments (NoAS, descriptor D14), the mean number of zero crossings per active segment (ZCAS, descriptor D16), the standard deviation of zero crossings per active segment (sdZCAS, descriptor D18), the mean number of local maxima per active segment (LocMaxAS, descriptor D15) and the standard deviation of local maxima per active segment (sdMaxAS, descriptor D17). To analyze the signal curve within an active segment, a method is used that interprets the output of the NLEO as a probability density

Table 1: Annotated CFAE.

CFAE-Class	MD A	MD B	Coinciding
C0	156	154	144
C1	107	143	84
C2	263	191	148
C3	79	117	53

MD, physician; A and B

function. The standard deviation of this function is a measure of the concentration of the signal curve in time direction. The mean of the standard deviation of all active segments per signal is the descriptor mean variance in time direction (MVarTD, descriptor D1).

Phase space descriptors: A combined view of the signal's magnitude and the alteration at a position t_i is enabled by the phase space. Here, the derivative of the signal $x'(t_i)$ is plotted over its magnitude $x(t_i)$. To weight the alteration and the magnitude equally, both are normalized to their maximum. To analyze the distribution of the samples in phase space, the phase space is divided into circular regions. Therefore a maximal distance d_{max} to the point of origin is defined. The distance is calculated by

$$d(t_i) = [x'(t_i)^2 + x(t_i)^2]^{1/2} \qquad (4)$$

To be more robust against outliers d_{max} is defined as the mean of the 5% of the maximal distances ($d_{0.05max}$)

$$d_{max} = \frac{1}{0.05N}\sum_{j \in s} d(t_j), \quad with\ S = \{i \mid d(t_i) \in d_{0.05max}\} \qquad (5)$$

The aim of this definition is to use enough samples to lower the influence of outliers to the outer boundary, but with a minimum of samples. The phase space is divided into four regions depending on the distance d_{max}. Region 1 encloses all sample values within a distance of 0.05 d_{max}. The border for region 2 is at 0.1 d_{max}, for region 3 at 0.2 d_{max}, and for region 4 at d_{max}. The number of samples in each region divided by the total number of samples is called "Phase Space Sample Ratio" (PSSR 1, PSSR 2, PSSR 3, PSSR 4) (descriptors D3–D6). Therefore, the PSSR will group periodic electrograms (e.g., sinus rhythm) usually into region 1. For chaotic or non-periodic signals, the assignment is more random. To reflect this behavior a binary function is generated. This binary function equals "1", if a sample is located in region 4 and "0" otherwise. As the entropy mirrors the information content of a random process, it is suitable to measure the periodicity of this binary function. The entropy of the binary function is termed as entropy of phase space (EPS 4, descriptor D7). To

evaluate the periodicity of the binary function, the standard deviation of distances between two adjacent rising edges is calculated. The resulting descriptor is called mean cycle length of phase space (MCPS 4, descriptor D8).

Wavelet based descriptors: Fractionated activity exhibits complex and time varying morphology. This is reflected by the broad frequency spectrum present in those electrograms. Using a wavelet based approach allows the analysis of different frequencies on multiple scales.

According to the algorithm presented in [17] a descriptor is set up, that counts the zero crossings in a given signal as a measure for the fractionation of this signal. Therefore, the signal is decomposed into coefficients with wavelet Coiflet 4 up to level 10. As shown by [19] Coiflet 4 is a suitable wavelet to analyze fractionation of electrograms. The algorithm can detect regions with high magnitude and high slew rate. As intracardiac electrograms often differ in terms of frequency, the wavelet level with the largest magnitude is searched. According to the algorithm from [17] this level is multiplied with the two preceding (higher frequency) levels. As a result, active regions will be emphasized and inactive regions will be suppressed. In the resulting signal, x_{res}, fractionated regions will be detected by a search for zero crossings. The number of zero crossings in this wavelet based electrograms are used as a descriptor (FracSig, descriptor D13).

Similarity of active segments: Faes et al. [10] described an algorithm that compares the similarity of different regions of an intracardiac electrogram. Inspired by Faes' work a similarity analysis is computed. The envelope of the absolute value of the analytical signal is calculated according to the algorithm described in [34]. The envelope follows the shape of the signal, but is always positive. This approach respects the increasing variation of the accurate shape in a more fractionated electrogram. Active segments are extracted using the NLEO-based segmentation algorithm. With these extracted segments a correlation matrix is built. First the segments are aligned using cross-correlation. On the overlapping parts, the absolute value of the correlation coefficient is calculated. On the basis of this correlation matrix, a clustering is performed. Starting with cluster of the two most similar segments, step by step the nearest similar segment is added and the similarity between the new segment and the cluster is computed. According to the definition of Kaufman and Rousseeuw [13] the similarity between two clusters is calculated from

$$s(R,Q)=\frac{1}{|R||Q|}\sum_{i\in R, j\in Q} s(i,j) \tag{6}$$

where $|R|$ and $|Q|$ are the cardinalities of both clusters, $s(i, j)$ is the similarity between element i from cluster R and element j from cluster Q. Here, the number of elements of cluster Q is set to 1. Finally, the mean value of the similarities of an active signal (Similarity AS, descriptor D9) is computed.

Amplitude statistics based descriptors: The histogram of an electrogram indicates the measured values during the time of recording. Little electrical activity in atrial electrograms is resulting in an amplitude around zero. So, the amplitude histogram has a high peak around "0", whereas the amplitude distribution of electrograms with strong electrical activity is spread more widely and is more gaussianlike. This behavior can be described with the fourth standardized moment, the kurtosis [25]. To lower the impact of outliers the signal is divided into k segments with a length l_s. The global kurtosis

(HistKurt, descriptor D2) is averaged over the kurtosis of the k segments. The segment length l_s is set to 1 s to make sure that even in sinus rhythm up to a heart frequency of 1 Hz there is at least one atrial excitation in this segment.

$$\text{HistKurt}=\frac{1}{k}\sum_{j=0}^{k-1}\text{kurtosis}\big(x(t\,|\,j\cdot l_s \leq t < (j+1)\cdot l_s)\big) \tag{7}$$

Table 2 gives an overview of the 18 descriptors, most of them used by the FDT for classification of the electrograms.

Classification/fuzzy decision tree

Building the fuzzy decision tree: As a major advantage, FDT avoids sharp split nodes but instead assigns objects in a region around the split point a relative affiliation to the following child nodes. This fuzzy zone is realized using sigmoid functions of form:

$$z_l(\omega)=1-\frac{1}{1+\exp\{-\sigma(v_\omega-s)\}} \tag{8}$$

$$z_r(\omega)=\frac{1}{1+\exp\{-\sigma(v_\omega-s)\}} \tag{9}$$

as proposed by Chandra and Varghese [7]. z_l and z_r are the affiliation to the left and the right child node, s is the split point and v_ω is the descriptor value of descriptor ω. The coefficient σ is the standard deviation of this descriptor on the training data. If the descriptor values are scaled with a factor a, the width of the fuzzy zone is changing by a^2. This is not intended.

In this decision tree, we defined a coefficient k to replace σ. The beginning z_{start} of the fuzzy zone is set when the sigmoid function equals 0.01 and the end z_{end} is set when the sigmoid function equals 0.99. In general $z_{start}(x)=g$ and $z_{end}(x)=1-g$. The range of a descriptor is the interval $[x_{min}, x_{max}]$. The width of the fuzzy zone is chosen as part $p\in[0, 1]$ of the interval length $i=x_{max}-x_{min}$. It is assumed that the fuzzy zone with width $p\cdot i$ is set symmetrically to the split point s. This leads to

$$z_{start}\left(s-\frac{p\cdot i}{2}\right)=g \quad \text{and} \quad z_{end}\left(s+\frac{p\cdot i}{2}\right)=1-g. \tag{10}$$

The sigmoid function is then expressed as

$$z(x)=\frac{1}{1+\exp\{-k(s-x)\}}. \tag{11}$$

Inserting (10) in (11) and solving for k results in

$$k_1=-\frac{\ln(g)-\ln(1-g)}{-(p\cdot i)/2}, \quad k_2=\frac{\ln(g)-\ln(1-g)}{(p\cdot i)/2}, \tag{12}$$

with $k_1=k_2$ one has

$$k=-\frac{2[\ln(1-g)-\ln(g)]}{p\cdot i} \tag{13}$$

The boundaries for the interval $[x_{min}, x_{max}]$ have to be determined from the training data. To be robust against outlier, the upper boundary value is set to 1.5 interquartile range of the upper quartile and the lower boundary to 1.5 interquartile range of the lower quartile. While

Table 2: Results of the mostly used descriptors.

Descriptor		CFAE 0		CFAE 1		CFAE 2		CFAE 3	
		$\tilde{\mu}$	iqr	$\tilde{\mu}$	iqr	$\tilde{\mu}$	iqr	$\tilde{\mu}$	iqr
D1	MVarTD	0.17	(0.01)	0.18	(0.02)	0.21	(0.02)	0.23	(0.03)
D2	HistKurt	22.62	(10.36)	12.88	(5.33)	9.76	(3.17)	6.83	(2.48)
D3	PSSR 1	0.88	(0.08)	0.70	(0.14)	0.52	(0.11)	0.35	(0.12)
D4	PSSR 2	0.02	(0.03)	0.07	(0.03)	0.13	(0,03)	0.14	(0.03)
D5	PSSR 3	0.01	(0.01)	0.04	(0.02)	0.08	(0.02)	0.11	(0.02)
D6	PSSR 4	0.10	(0.03)	0.18	(0.09)	0.26	(0.08)	0.40	(0.13)
D7	EPS 4	1.60	(0.19)	2.09	(0.44)	2.51	(0.26)	2.67	(0.08)
D8	MCPS 4	5.71	(0.32)	5.14	(0.31)	4.74	(0.20)	4.59	(0.14)
D9	Similarity AS	-0.17	(0.42)	-0.68	(0.26)	-0.97	(0.17)	-1.17	(0.21)
D10	AR	0.18	(0.06)	0.32	(0.15)	0.56	(0.17)	0.82	(0.14)
D11	MLAS [ms]	54.46	(13.42)	75.67	(31.97)	132.09	(62.60)	312.88	(271.54)
D12	sclMLAS [ms]	4.45	(13.95)	18.01	(13.58)	86.25	(78.23)	345.59	(334.66)
D13	FracSig	345.00	(112.00)	630.00	(273.50)	809.00	(265.00)	1152.00	(433.00)
D14	NoAS	18.00	(5.00)	26.00	(5.00)	24.00	(4.00)	15.00	(9.25)
D15	LocMaxAS	1.68	(0.29)	2.02	(0.41)	2.52	(0.40)	3.29	(0.78)
D16	ZCAS	1.44	(0.34)	1.76	(0.42)	2.27	(0.41)	3.06	(0.76)
D17	sdMaxAS	0.18	(1.76)	1.48	(1.01)	4.21	(1.69)	6.54	(2.03)
D18	sdZCAS	-0.13	(1.78)	1.40	(1.01)	3.68	(1.70)	6.07	(2.09)

Median ($\tilde{\mu}$) and interquartile range (iqr) per CFAE class are given.

training the decision tree, the aim is to find a split point at a node t, so that the relative membership of the objects to the right class is increasing. The Gini diversity index (GDI) is a regular criterion to evaluate possible split points [3]. The adaption of the GDI for usage in a FDT was made by Chandra et al. [7]:

$$GDI(s_j) = \sum_{v=1}^{2} \frac{N^{(t_v)}}{N^{(t)}} \left[1 - \sum_{j=1}^{J} \left(\frac{N_j^{(t_v)}}{N^{(t_v)}} \right)^2 \right] \tag{14}$$

where J is the number of classes; $N^{(t_v)}$ is sum of the fuzzy-membership values of records of child node t_v with chosen split point s_j on descriptor j. $N^{(t)}$ is the sum of the fuzzy-membership values of records in the tth partition before split. $N_j^{(t_v)}$ is the sum of the product of fuzzy-membership values of the attribute and the fuzzy-membership values of the corresponding record for class j. The GDI is a rating criterion to find possible split candidates. To find those split points for a descriptor vector we do the following:

– sort descriptor values in descending order,
– set the positions where the class affiliation is changing as candidates for a split point,
– set the split point to the arithmetic mean of both neighboring descriptor values.

According to this algorithm for each descriptor vector the split point candidates are found. Finally, the candidate with the smallest GDI is chosen. The width of the fuzzy zone can be varied according to the distribution of classes over a descriptor. If there is a large overlap, the fuzzy zone should be larger than in the case of separated classes.

In this work, the best width of the fuzzy zone was determined by choosing the best result from a set of different fuzzy zones (Figure 2). For each fuzzy zone a 10×10 cross validation was realized and the mean error rate for each fuzzy zone was computed. The best error rate was achieved for a fuzzy zone with a width of 20%.

To make a decision on the class belonging of a given record, the decision tree needs an inference instruction. In contrast to a decision tree with sharp split values, where a record will get the class membership of the leaf node to which it is assigned, here, a test record can be assigned to more than one leaf node by a FDT. The inference instruction for a given record X_{Test} considers the fuzzy membership to a leaf node $z(t_i, X_{Test})$ and its class membership $c(t_i)$. The fuzzy membership of a leaf node results from the product of fuzzy memberships of the passed nodes. The total class membership of a record X_{Test} is defined as

$$c_{total} = (X_{Test}) = \sum_i c(t_i) \cdot z(t_i, X_{Test}) \tag{15}$$

The classification of a record results in partial affiliation to the CFAE classes. From these partial affiliations, the resulting CFAE class is derived by a majority decision. The classified record will get the CFAE class label of the class with the highest percentage. The percentage can also be shown as a measure of certainty.
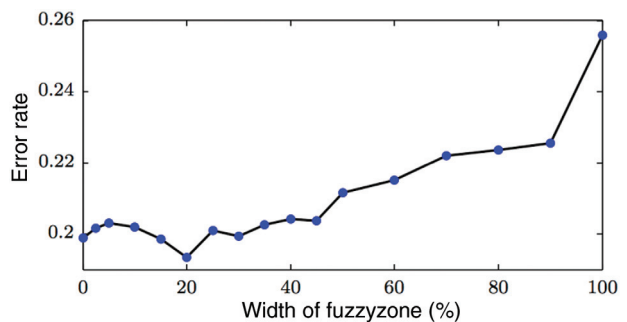

Figure 2: Evaluation of different fuzzy zone widths to find the "best" fuzzy zone width. Displayed is the error rate as result of a 10×10 cross validation for each width.

Chandra and Varghese used different sharp stop criteria to limit the growth of the tree while training [7]. This technique has some limitations. Therefore, the stop criteria of this decision tree were given very weak definitions. This led to an oversized tree at first. In an additional step this oversized tree was optimized to the best size [3]. The aim of pruning was to optimize the tree in regard to the complexity and the predictive accuracy. A sequence of trees is produced by the use of cost-complexity-pruning [3]. Out of this sequence the best tree is chosen by evaluating the error rate in consideration with the size of the FDT. The final tree for the classification in a clinical context will be generated on the whole set of available classification data. The best tree is chosen again after cross validation and optimization by evaluating the size and error rate.

Validation and statistical analysis

The standard way to evaluate a classificator is to perform a 10×10 cross validation [3]. This means the data set has to be separated into 10 equal subsets. The tree is trained with nine subsets and tested with the remaining "unknown" subset. The correct rates and error rates are computed and rated. This "training, testing and evaluation" step has to be repeated until each of the 10 subsets was used to test the tree. This whole procedure is repeated 10 times with different randomly created subsets, which leads to a total of 100 trees that are evaluated.

To optimize this tree, the evaluation has not only been done on the CFAE database but also on datasets from the UCI Machine Learning Database [22]. The UCI database is a regular database to test and evaluate machine learning algorithms. From this database five data sets [Iris, Wine, Wisconsin Breast Cancer (WBC), Haberman's Survival (HS), and Glas Identification (GI)] have been chosen to evaluate the FDT. Data sets with continuous feature values and multi-class assignments have been chosen to be most similar to the CFAE data. Finally the results have been compared to different decision tree algorithms (CART [3], SLIQ [23], C4.5 [35], FDTx [7]).

Evaluation of clinical data

To evaluate the distribution of the CFAE classes in clinical cases, bipolar electrograms were recorded for 5 s with a 300 Hz filter setting. Anatomical and electrical information were exported from the system and the 3D anatomical shells as well as the corresponding electrograms were reconstructed. At first, the algorithm was set to detect and display the mean cycle length (CL) of AF (mean dV/dt), with a deflection width of 10 ms and a refractory period of 37 ms. Electrograms with a mean CL>120 ms were color coded in violet, 120–70 ms in rainbow colors and <70 ms in white. The resulting map was compared to the CFAE classification map using the FDT.

Results

Descriptors

All descriptors have been calculated using the database described previously. In the training phase, the FDT chooses the best descriptors to separate the classes depending on the cost function (e.g., GDI). The results of the 18 most frequently used descriptors are shown in Table 2.

For each CFAE class the median ($\tilde{\mu}$) and the interquartile range (*iqr*) per descriptor is shown. The $\tilde{\mu}$ and the *iqr* give an overview of the distribution of the descriptor values and the power to separate CFAE classes of each descriptor. Generally, there is no descriptor that can separate between all four CFAE classes alone. However, the descriptors can separate between different CFAE class groups. For example the descriptors PSSR 3 (D5) and AR (D10) can separate C0 from C3. The mean NoAS (D14) can be used to separate C0 and C3 from C1 and C2. The boxplots of these descriptors are depicted in Figure 3. To evaluate the results of the phase space descriptors, the numbers of samples per region are divided by the total number of samples per signal. The boxplot of the PSSR 3 (D5) is depicted in Figure 3. As can be seen, the median values are significantly different on a 5% level (boxplot notches) for all four CFAE classes. Also there is no overlapping of the *iqr*. The same applies to the AR. The range for the AR varies between 0 (no activity) and 1 (continuous activity). It increases from 0.18 (C0) to 0.32 (C1) and from 0.56 (C2) to 0.82 (C3). An increasing value is correlated therefore with an increase in entropy of the electrogram. The entropy of the number of hits in the different regions of the phase space (EPS 4) is an often used descriptor, too. EPS 4 can be used to distinguish between C0 and C2 and C3. The median values for NoAS (D14) are 18 (C0), 26 (C1), 24 (C2), and 15 (C3), respectively. The values are increasing from CFAE class C0 to class C1. The frequency and fractionation is also increasing from C0 to C1. For C1 and C2, the NoAS are nearly the same as these two states are similar with regard to the number of CFAE. For C3 the activity is getting more continuous and therefore the NoAS (D14) is decreasing again. In the case of continuous activity it will be one. Finally, the NoAS can separate C0 and C3 from C1 and C2. There is an overlap in the *iqr* of C1 and C2, but, although the value range is very similar the $\tilde{\mu}$ are significantly different. For this four delineated descriptors the $\tilde{\mu}$ are significantly different on a 5% level for all CFAE classes.

On the one hand, with increasing atrial activity the signals have more deflections and zero crossings and the iso-electric line is vanishing. Hence, the amplitude histogram is broad and the HistKurt (D2) has small or negative values. Sinus rhythm electrograms, on the other hand, mainly present an iso-electric line with some deflection other than zero. Therefore, the amplitude histogram has a pronounced peak. In this case the kurtosis values will be larger than for the ones for the continuous activity. This behavior is reflected by descriptors D1 and D2. When the
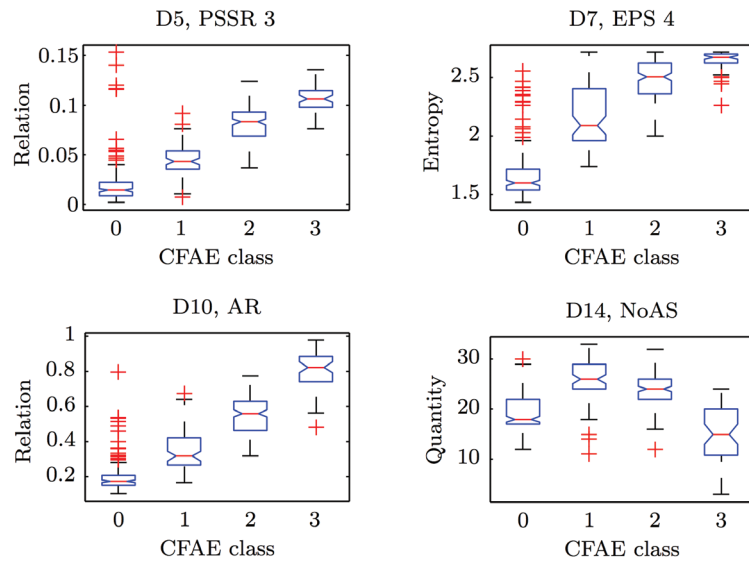
Figure 3: Boxplots of four descriptors used by the decision tree depicted in Figure 4. Red crosses mark outliers.

heart rhythm evolves from sinus rhythm to AF the similarity of activation patterns gets lost [37]. As a consequence the active segments are getting dissimilar (D9).

Class C0 and C1 electrograms contain mostly regions without alteration (iso-electric line is prominent). This type of signals will be represented primarily in the phase space in region 1 and 2, whereas CFAE class C2 and C3 will be more prominent in region 3 and 4. Also the alterations of C2 and C3 signals make them lie more in the outer limits of the phase space. Descriptors D3–D8 mirror this. Going from the origin to the boundary of the phase space the number of C0 samples is decreasing while the number of C3 samples are increasing. PSSR 3 (region 3) contains 1% of C0, 4% of C1, 8% of C2, and 11% of C3 signals. EPS 4 and MCPS 4 give an information of the presence of a CFAE class in phase space region 4. These descriptors reflect the same outcome as the descriptor PSSR. The descriptors D13, D15, D16, D17, and D18, which reflect the fractionation

of the signal, are increasing from C0 to C3, too. This corresponds to the increasing atrial activity and the increasing number of active fractionated segments (Figure 3).

Validation of the tree implementation

The FDT presented in this work was compared to the results of decision trees with sharp split values and the FDT by Chandra and Varghese (FDT_x) [7]. Table 3 gives an overview of the results. Presented is the error rate and standard deviation for each classifier after a 10×10 cross validation. The results for FDT_x and SLIQ were obtained from [7] and the results for CART and C4.5 were obtained from [31]. The decision trees CART, SLIQ, and C4.5 have sharp split values; the FDT by Chandra and Varghese uses a sigmoidal function as decision border. FDT is the tree presented in this work. In addition to the error rate, the fuzzy zone width for each

Table 3: Comparison of different decision tree algorithms.

	CART	SLIQ	C4.5	FDT_x	FDT	Fuzzy zone width
Iris	93.5±0.8	98.0±3.2	95.1±0.6	98.0±3.2	96.1±4.8	0.10
Wine	89.3±0.8	88.3±8.1	92.7±1.1	88.9±4.5	89.1±7.2	0.01
WBC	–	93.3±4.8	–	96.4±4.9	92.4±3.6	0.20
HS	–	65.8±10.8	–	72.6±7.8	74.4±7.2	0.05
GI	67.7±1.6	65.0±16.0	68.6±2.0	68.6±8.7	69.4±8.7	0.10

Shown are the error rate (%)±the standard deviation (%) after 10×10 cross validation. CART and C4.5 data are obtained from Olaru and Wehenkel [31], SLIQ and FDTx are obtained from Chandra and Varghese [7]. Wisconsin Breast Cancer (WBC), Haberman's Survival (HS), Glas Identification (GI).

Table 4: Wrong assignments per CFAE class in %.

	CFAE 0	CFAE 1	CFAE 2	CFAE 3
CFAE 0	X	72	24	4
CFAE 1	37	X	63	1
CFAE 2	2	49	X	49
CFAE 3	0	0	100	X

data set is given. When comparing with other classifiers for the data sets Iris, Wine and WBC, the proposed FDT delivers comparable results. Focusing on the results of the Haberman's Survival and the Glas Identification data sets, the FDT is significantly superior.

Resulting fuzzy decision tree

Applying a 10×10 cross validation on the presented FDT with a fuzzy zone width of 20% results in a mean correct rate of 80.65±3.3%. Correct rates for the different CFAE classes are C0: 83.1±4.4%, C1: 81.0±8.1%, C2: 75.8±8.4%, C3: 82.7±17.8%. The distribution of the wrong assignments of signals per class is shown in Table 4. Increasing the training data will lead to an increased accuracy of the FDT. The depicted tree in Figure 4 is chosen from the sequence of cross validated trees. This tree has a correct rate of 86.1% (Figure 4).

Visualization of CFAE classes

Using the CFAE mean algorithm, areas with high atrial frequencies were located in different areas of the left atrium. A loose correlation between CFAE class 2 and 3 could be estimated. Interestingly, in the FDT map electrograms with continuous activity (C3) were surrounded by C2 and C1. C3 were present mainly around the PV ostia, the anterior wall and the inferior part of the posterior wall (Figure 5).

Discussion

The ablation of persistent AF is still challenging. Up to date, intracardiac electrograms are the only information which can be obtained during the procedure. But these electrograms present a broad variation. Current detection algorithms for CFAE are mainly based on single characteristics but do not respect the full complexity of the electrograms. Therefore, it seems appropriate to use more than one descriptor. Based on the modified Wells' criteria, we classified CFAE into four subgroups with a continuously increase in complexity (C0–C3). The FDT can be used to classify electrograms which present a progressive transition between the different classes. This overcomes the results sensitivity/vulnerability against small changes of
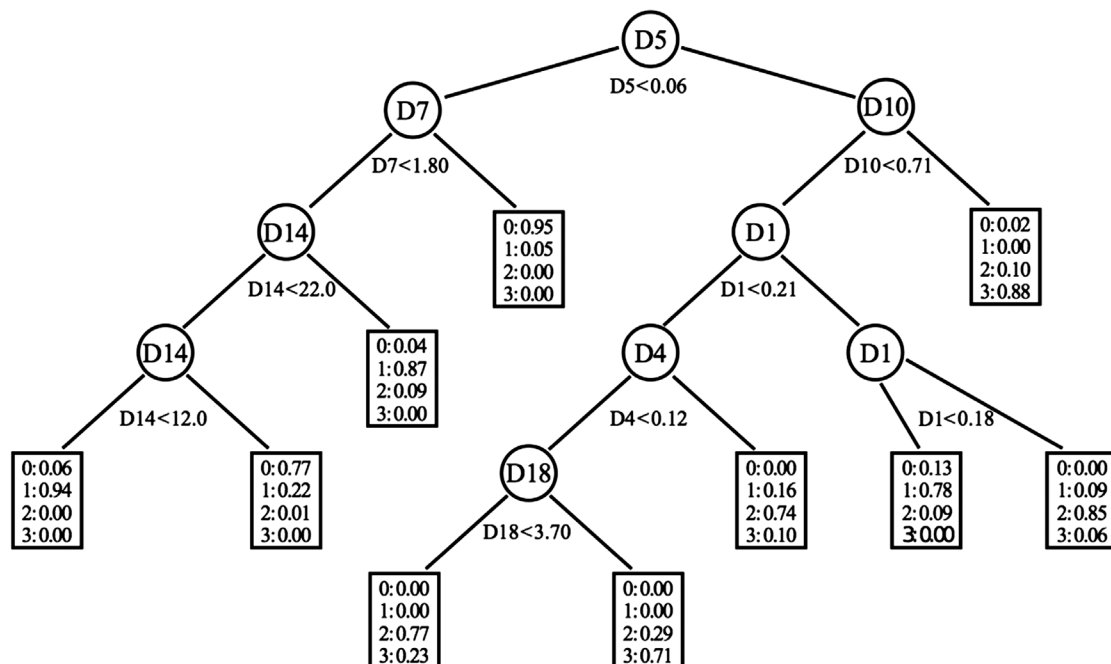
Figure 4: Example of the optimal tree chosen from the cross validation process.
At each node the chosen descriptor and its split value is depicted. For each leaf node the membership result is shown. The correct rate for this tree is 86.1%.
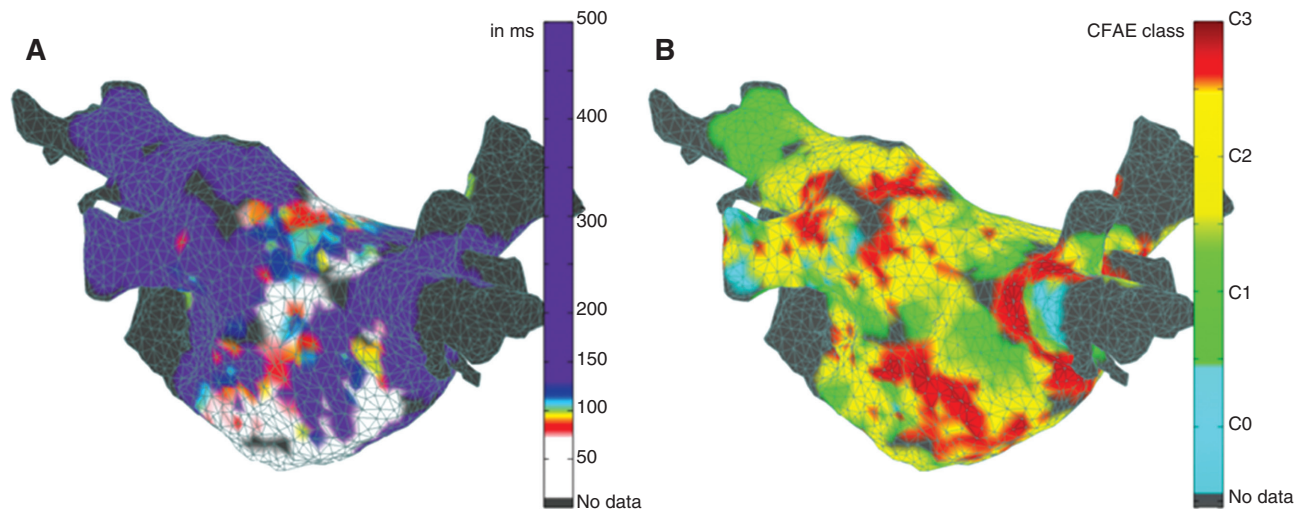
A

B

Figure 5: Example of a clinical evaluation of the CFAE classes using the FDT in comparison to a CFAE mean map.
View of the posterior wall with left upper and lower pulmonary vein (left) and right upper and lower pulmonary vein (PV) (right).
(A) automated CFAE mapping algorithm (Ensite NavX). Bipolar electrograms were recorded for 5 s, filter settings were at 30–500 Hz. The algorithm was set to detect and display the mean CL of atrial fibrillation (mean dV/dt), with a deflection width of 10 ms and a refractory period of 37 ms. Electrograms with a mean CL>120 ms were color coded in violet, 120–70 ms in rainbow colors and <70 ms in white.
(B) CFAE classes using the FDT. C0 (blue): non-fractionated atrial electrogram of high frequency, C1 (green): fractionated atrial electrogram with periodic activity, C2 (yellow): mixture of periodic fractionated and periodic non-fractionated atrial electrograms; C3 (red): electrogram with continuous activity. Electrograms with continuous activity C3 are surrounded by C2 and C1. C3 were present mainly around the PV ostia and the inferior part of the posterior wall.

the descriptor's value, which is one of the main shortcomings of decision trees with sharp split points [7]. Instead of using sharp split points to separate classes, the FDT assigns test objects within the fuzzy zone to classes with a specific probability. In this tree, several descriptors are introduced to describe the different characteristics of CFAE. The tree classifies a given signal into one of the four groups and presents a percentage of certainty.

Teager's NLEO was used to differentiate between active and passive segments of the electrograms with simultaneous consideration of frequency and amplitude. The MLAS (D11), sdMLAS (D12) as well as AR (D10) are increasing from C0–C3 which reflects a faster and more complex local activation. These algorithms are therefore able to distinguish between no activity and continuous activity. The NoAS (D14) can separate C0 and C3 from C1 and C2 and the EPS (D7) between C0 and C3. PSSR 3 (D5) and the AR (D10) are able to distinguish between C1 and C2.

In the training phase, 18 descriptors in 100 different trees were analyzed and the correct rate of the CFAE classification was analyzed.

The performance of the FDT improves the more data is used to train the tree. In this study 429 signals with coinciding interpretation were analyzed. The smallest group

of CFAE classes was C3 with 53 electrograms. Therefore, a maximum of 53 electrograms could be used in each CFAE class for the cross validation. To overcome the problem that a 10×10 cross validation would lead into a statistical analysis with only 5 signals, we performed a 10×5 cross validation. Using this technique, 11 signals remained to test the tree.

The width of the fuzzy zone for the FDT was set to 20%. Although another model, based on the GDI for different kind of class contribution (normally and unequally distributed, different overlap regions), revealed small fuzzy regions. In our experience, an increased fuzzy zone improves the power of the FDT (Figure 2).

The evaluation by the cross validation process estimates the correct rate of the method. This value is correct whilst the training data set reflects reality data. To generate a FDT working with a new data set or in a clinical setting, the tree will be trained with all available data [41]. The correct rate for this new classifier can be estimated from the mean correct rate of the cross validation. Therefore, a mean correct rate of 81±3% can be expected. The depicted tree in Figure 4 is chosen from the sequence of cross validated trees. This tree has a correct rate of 86.1%. This reflects the possibility, that with a wider set of training data, the outcome of the FDT can still be improved. In

addition to the classification results, a percentage of certainty is given for each electrogram.

Compared to other algorithms, the proposed tree presents similar correct rates for Iris, Wine, and WBC data sets. For the Haberman's Survival and Glas Identification data sets the proposed FDT is significantly superior. The reason for that might be the more fuzzy features of these data sets. This implies the strength of the proposed tree on fuzzy data.

In clinical use, the algorithm is able to display four different CFAE subgroups. This enables a more transparent and objective way of CFAE interpretation. For CFAE ablation, the most promising electrograms present continuous activity. These electrograms can be detected in 100% of the cases and accentuated in the 3D map. However, the classification of the CFAE allows a more detailed evaluation and may improve the knowledge about the stability and instability of certain areas during AF. This may restrict the search for characteristic electrograms sustaining AF.

Conclusion

CFAE express an important element of the AF substrate. We propose a new algorithm for automatic CFAE classification. According to the modified classification by Wells et al., CFAE were divided into four subgroups [42]. The subgroups were defined as non-fractionated with high frequency, fractionated with periodic activity, instable electrograms with a mixture of periodic fractionated and periodic non-fractionated atrial electrograms, and continuous activity. The algorithm is based on a FDT including 18 descriptors. Given a set of training data, a FDT classifier can be constructed. The training algorithm automatically chooses the ideal combination of descriptors and split points to the classification of training data. Obviously, this can lead to a tree, which only uses a subset of the offered set of descriptors as shown in the example tree (Figure 4). Using this tree, CFAE were sorted to one of the subgroups with a correct rate of at least 81%±3%. Electrograms with continuous activity were detected correctly 100% of the time. In addition, a percentage of certainty is given for each electrogram. The FDT is therefore able to classify CFAE with respect to their progressive transition. This will objectify CFAE interpretation in a more reproducible and transparent way.

Limitations

This is a retrospective analysis of electrograms recorded from patients with AF. Prospective studies are needed to evaluate the clinical impact of this algorithm.

References

[1] Allessie MA, Konings K, Kirchhof CJ, Wijffels M. Electrophysiologic mechanisms of perpetuation of atrial fibrillation. Am J Cardiol 1996; 77: 10A–23A.

[2] Barbaro V, Bartolini P, Calcagnini G, Morelli S, Michelucci A, Gensini G. Automated classification of human atrial fibrillation from intraatrial electrograms. Pacing Clin Electrophysiol 2000; 23: 192–203.

[3] Breiman L, Friedman JH, Olshen RA, Stone CJ. Classification and regression trees. New York, NY: Chapman & Hall, 1984.

[4] Brooks AG, Stiles MK, Laborderie J, et al. Outcomes of long-standing persistent atrial fibrillation ablation: a systematic review. Heart Rhythm 2010; 7: 835–46.

[5] Calkins H, Kuck KH, Cappato R, et al. 2012 HRS/EHRA/ECAS expert consensus statement on catheter and surgical ablation of atrial fibrillation: recommendations for patient selection, procedural techniques, patient management and follow-up, definitions, endpoints, and research trial design. Europace 2012; 528–606.

[6] Camm J, Lip GYH, De Caterina R, et al. 2012 focused update of the ESC Guidelines for the management of atrial fibrillation: An update of the 2010 ESC Guidelines for the management of atrial fibrillation * Developed with the special contribution of the European Heart Rhythm Association. Eur Heart J 2012.

[7] Chandra B, Varghese PP. Fuzzy SLIQ decision tree algorithm. IEEE Trans Syst Man Cybern B Cybern 2008; 38: 1294–1301.

[8] Ciaccio EJ, Biviano AB, Whang W, Garan H. Identification of recurring patterns in fractionated atrial electrograms using new transform coefficients. Biomed Eng Online 2012; 11: 4.

[9] Cuesta-Frau D, Cirugeda-Roldan E, Pico AM, Novak D, Kremen V. Atrial electrogram complex fractionated entropy study. Exp Clin Cardiol 2014; 20: 5566–5574.

[10] Faes L, Nollo G, Kirchner M, et al. Principal component analysis and cluster analysis for measuring the local organisation of human atrial fibrillation. Med Biol Eng Comput 2001; 39: 656–663.

[11] Go AS, Hylek EM, Phillips KA, et al. Prevalence of ' implications for rhythmmanagement and stroke prevention: the AnTicoagulation and risk factors in atrial fibrillation (ATRIA) study. J Am Med Assoc 2001; 285: 2370–2375.

[12] Haïssaguerre M, Jaïs P, Shah DC, et al. Spontaneous initiation of atrial fibrillation by ectopic beats originating in the pulmonary veins. N Engl J Med 1998; 339: 659–666.

[13] Kaufman L, Rousseeuw PJ. Finding groups in data an introduction to cluster analysis. Wiley Series. New York: Wiley Interscience 1990.

[14] Keller MW, Schuler S, Wilhelms M, et al. Characterization of radiofrequency ablation lesion development based on simulated and measured intracardiac electrograms. IEEE Trans Biomed Eng 2014; 61: 2467–2478.

[15] Khawaja A. Automatic ECG analysis using principle component analysis and wavelet transformation. Univ Karlsruhe 2007; 3: 1–2.

[16] Khawaja A, Sanyal A, Doessel O. A wavelet-based multi-channel ECG delineator. In: 3rd Eur Med Biol Eng Conf; 2005; Vol.11.

[17] Kim KH, Kim SJ. A Wavelet-based method for action potential detection from extracellular neural signal recording with low signal-to-noise ratio. IEEE Trans Biomed Eng 2003; 50: 999–1011.

[18] Kordik P. Fully automated knowledge extraction using group of adaptive models evolution. Czech Tech Univ Prague Fac Electr Eng 2006; (September): 1–150.

[19] Křemen V, Lhotská L, Macaš M, et al. A new approach to automated assessment of fractionation of endocardial electrograms during atrial fibrillation. Physiol Meas 2008; 29: 1371–1381.

[20] Lau DH, Maesen B, Zeemering S, Verheule S, Crijns HJ, Schotten U. Stability of complex fractionated atrial electrograms: a systematic review. J Cardiovasc Electrophysiol 2012; 23: 980–987.

[21] Lee G, Roberts-Thomson K, Madry A, et al. Relationship among complex signals, short cycle length activity, and dominant frequency in patients with long-lasting persistent AF: a high-density epicardial mapping study in humans. Heart Rhythm 2011; 8: 1714–1719.

[22] Lichman M. UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science 2013.

[23] Mehta M, Agrawal R, Rissanen J. SLIQ: A fast scalable classifier for data mining. In: Apers P, Bouzeghoub M, Gardarin G, editors. Adv database technol – EDBT '96 SE – 2. vol. 1057. Berlin Heidelberg: Springer 1996: 18–32.

[24] Nademanee K, McKenzie J, Kosar E, et al. A new approach for catheter ablation of atrial fibrillation: mapping of the electrophysiologic substrate. J Am Coll Cardiol 2004; 43: 2044–2053.

[25] Najim K, Ikonen E, Daoud A-K. Stochastic processes: estimation, optimisation and analysis. London: Elsevier 2004.

[26] Navoret N, Jacquir S, Laurent G, Binczak S. Detection of complex fractionated atrial electrograms using recurrence quantification analysis. IEEE Trans Biomed Eng 2013; 60: 1975–1982.

[27] Ndrepepa G, Karch MR, Schneider MAE, et al. Characterization of paroxysmal and persistent atrial fibrillation in the human left atrium during initiation and sustained episodes. J Cardiovasc Electrophysiol 2002; 13: 525–532.

[28] Nguyen MP, Schilling C, Dossel O. A new approach for frequency analysis of complex fractionated atrial electrograms. *Eng Med Biol Soc 2009 EMBC 2009 Annu Int Conf IEEE* 2009; 368–371.

[29] Nielsen JC, Englund A. A randomized multicenter comparison of radiofrequency ablation and antiarrhythmic drug therapy as first line treatment in 294 patients with paroxysmal atrial fibrillation MANTRA-PAF investigators. N Engl J Med 2012; 367: 1587–1595.

[30] Nollo G, Marconcini M, Faes L, Bovolo F, Ravelli F, Bruzzone L. An automatic system for the analysis and classification of human atrial fibrillation patterns from intracardiac electrograms. IEEE Trans Biomed Eng 2008; 55: 2275–2285.

[31] Olaru C, Wehenkel L. A complete fuzzy decision tree technique. Fuzzy Sets Syst 2003; 138: 221–254.

[32] Oral H, Chugh A, Good E, et al. Radiofrequency catheter ablation of chronic atrial fibrillation guided by complex electrograms. Circulation 2007; 115: 2606–2612.

[33] Oral H, Chugh A, Yoshida K, et al. A randomized assessment of the incremental role of ablation of complex fractionated atrial electrograms after antral pulmonary vein isolation for long-lasting persistent atrial fibrillation. J Am Coll Cardiol 2009; 53: 782–789.

[34] Potamianos A, Maragos P. A comparison of the energy operator and the Hilbert transform approach to signal and speech demodulation. Signal Process 1994; 37: 95–120.

[35] Quinlan JR. Validation and statistical analysis. C4.5: Programs for machine learning. Revised, U. vol 5. Burlington, MA, USA: Morgan Kaufman Publ Inc., 1993: 1–302.

[36] Roberts-Thomson KC, Kistler PM, et al. Fractionated atrial electrograms during sinus rhythm: relationship to age, voltage, and conduction velocity. *Heart Rhythm* 2009; 6: 587–591.

[37] Schilling C, Luik A, Schmitt C, Doessel O. Analysis of intracardiac ECG measured in the coronary sinus. In: 4th Eur Congr Med Biomed Eng; 2009; 22: 260–263.

[38] Schilling C, Nguyen MP, Luik A, Schmitt C, Doessel O. Non-linear energy operator for the analysis of intracardiac electrograms. In: IFMBE Proc World Congr Med Phys Biomed Eng 2009; 25: 872–875.

[39] Schmitt C, Ndrepepa G, Weber S, et al. Biatrial Multisite Mapping of Atrial of Atrial Fibrillation. Am J Cardiol 2002; 89: 1381–1387.

[40] Smelley MP, Knight BP. Approaches to catheter ablation of persistent atrial fibrillation. Hear Rhythm 2009; 6: S33–S38.

[41] Varma S, Simon R. Bias in error estimation when using cross-validation for model selection. BMC Bioinformatics 2006; 7: 91.

[42] Wells JL, Karp RB, Kouchoukos NT, Maclean WAH, James TN, Waldo AL. Characterization of atrial fibrillation in man: studies following open heart surgery. Pacing Clin Electrophysiol 1978; 1: 426–439.

[43] Zrenner B, Ndrepepa G, Karch MR, et al. Electrophysiologic characteristics of paroxysmal and chronic atrial fibrillation in human right atrium. J Am Coll Cardiol 2001; 38: 1143–1149.