Classification of Atrial Tachycardia Types Using Dimensional Transforms of ECG Signals and Machine Learning

Samuel Ruipérez-Campillo[1,2,3,4], José Millet[2], Francisco Castells[2]

[1] Bioengineering Department, University of California in Berkeley, CA, USA
[2] ITACA Institute, Universitat Politecnica de Valencia, Valencia, Spain
[3] School of Medicine, Stanford University, Palo Alto, CA, USA
[4] D-ITET, Swiss Federal Institute of Technology (ETHz), Zürich, Switzerland

Abstract

Accurate non-invasive diagnoses in the context of cardiac diseases are problems that hitherto remain unresolved. We propose an unsupervised classification of atrial flutter (AFL) using dimensional transforms of ECG signals in high dimensional vector spaces. A mathematical model is used to generate synthetic signals based on clinical AFL signals, and hierarchical clustering analysis and novel machine learning (ML) methods are designed for the unsupervised classification. Metrics and accuracy parameters are created to assess the performance of the model, proving the power of this novel approach for the diagnosis of AFL from ECG using innovative AI algorithms.

1. Introduction

Machine Learning (ML) is leading the paradigm shift in the way we analyse and process cardiac signals for prediction of diseases [1] and physiological events [2]. Unsupervised learning is being used to identify parameters of the ECGs [3] [4]. Simulations, which are becoming more sophisticated for diseases such as atrial flutter (AFL) [5], could be of great use to train artificial intelligence (AI) algorithms.

We propose in this study a methodology based on signal processing techniques, ML, and mathematical modelling, to characterise and diagnose AFL, which has an increasing prevalence [6] and whose treatment could be more accurately planned with prior information. To that end, we establish a mathematical framework in the context of Hilbert spaces to represent our surface signals through dimensional transformations. We design metrics and agglomerative nesting techniques to cluster AFL groups in an unsupervised manner, providing an accurate diagnosis without any manual annotation.

2. Materials and Methods

2.1. Materials

In this study we consider slow velocity conduction regions as the main discriminant factor to classify different types of macroreentrant atrial tachyarrhythmias [7]. Thus, a variant of a synthetic model based on this principle [7] [8] is used to create 8,000 AFL VCG loops, of 8 different groups, which correspond to 4 different regions of slow conduction, and two propagation directions – clockwise (CW) and counterclockwise (CCW).

2.2. Defining Vector Space for VCGs

Let $\mathscr{J}$ represent a distribution in the context of Hilbert spaces and $\mathbb{F}$ a field, avoiding in our study any field out of the sets of real or complex numbers (i.e. $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$).

Any multichannel signal can be represented from the concatenation of its waveforms in the different channels as a vector in $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$ in a Hilbert Space framework. Each of the temporal samples of the concatenated waves is translated as the magnitude of each of the components of than inner-product space. This vector is a generalization of $\mathbb{R}^n$ of real (or complex) n-tuples, and is then analogous to the particular waveform (from the concatenated channels) of each patient (see figure 1).

2.3. Projecting the Signals

With the aim of working in a computationally efficient framework that can be generalised to more complex and demanding analyses, we propose an algorithm of dimensionality reduction based indistinctly on Singular Value Decomposition (SVD) or Principal Components Analysis (PCA). The resultant vector subspace will host the projected vectors of the original representation of signals in a Hilbert Space with fewer dimensions. In particular, for a space $\boldsymbol{S}$, every *Cauchy Sequence* of VCG samples as elements of that space must converge to an element of such space. In fact, for any positive $\lambda_i \in \mathbb{R}$, an inner product is a function such that $\boldsymbol{S}$ x $\boldsymbol{S} \rightarrow \mathbb{F}$: $(v, w) \longmapsto \langle v, w \rangle$, $\forall\, v, w \in \boldsymbol{S}$.

To that end, let us define $\boldsymbol{S}$ as the original M-dimensional space where the observations were firstly presented from the concatenated multichannel representation. We then define a non-empty subset of $\boldsymbol{S}$ and name it $\boldsymbol{R}$ over $\mathbb{F}^P$.

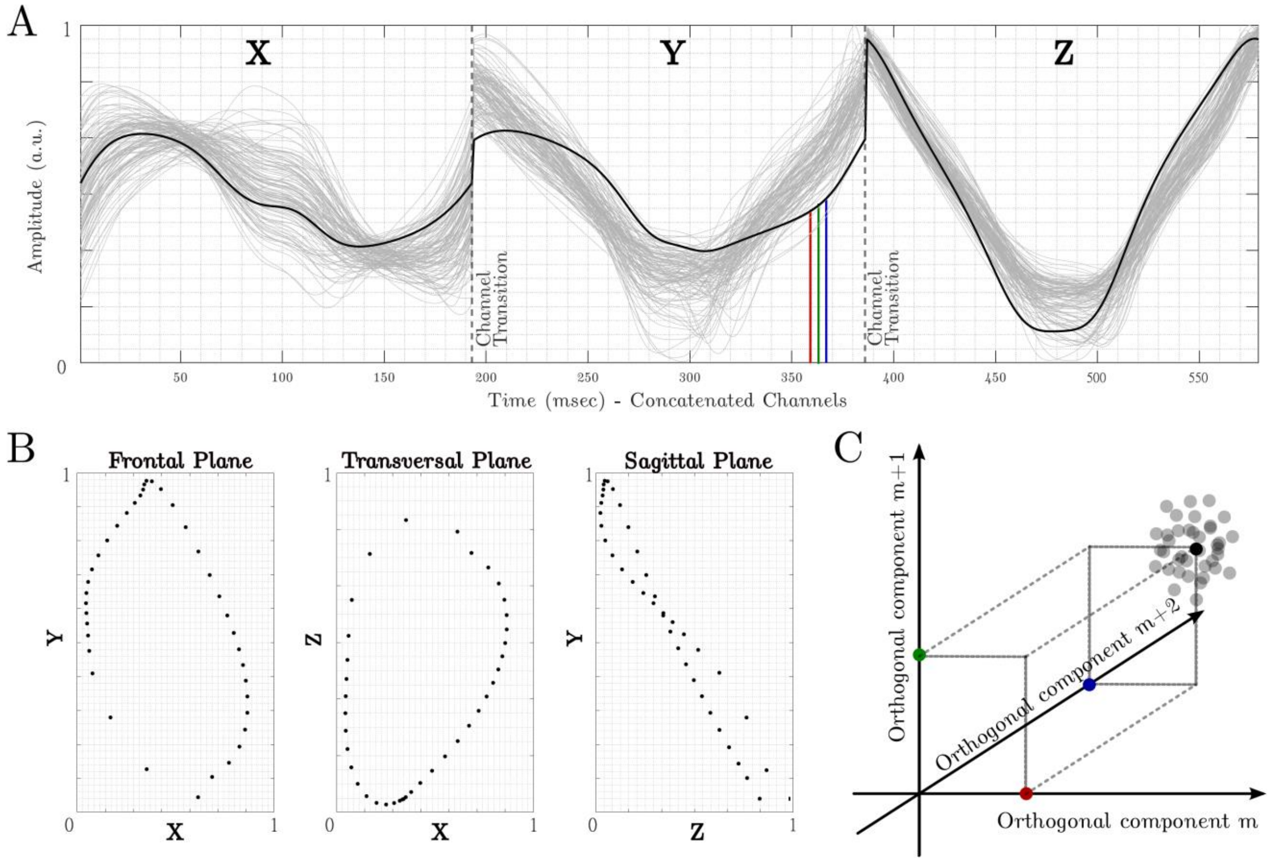

Figure 1. A: *Concatenation of the three channels for a synthetic type of AFL. Each of these samples is used to provide magnitude to a dimension of the n-dimensional vector space where each wave is represented as a point. An example for three samples is depicted in blue, green and red (see* **C***). One VCG is marked with a black line for illustration purposes (see* **B***).*

The necessary condition that needs to be fulfilled is that $\forall s \in \boldsymbol{S}$ and $f \in \mathbb{F}$ the following is true: $f \cdot s \in \boldsymbol{R}$; $\forall\ s, s' \in \boldsymbol{S}$; and $s + s' \in \boldsymbol{S}$. Hence, for a subset $\boldsymbol{V} = \{\upsilon_1, \upsilon_2, ..., \upsilon_N\} \subset S$ over the vector space $\boldsymbol{S}$ over $\mathbb{F}$, the VCG signals are represented in a space spanned by $\boldsymbol{V}$ and can be written as:

$$\mathrm{span}(\upsilon_1, \upsilon_2, \upsilon_3, \dots, \upsilon_N) := \left\{ \sum\nolimits_{\kappa} \xi_\kappa \upsilon_\kappa : \xi_\kappa \in \mathcal{F} \right\}$$

2.4. Dimensionality Reduction

The aforementioned projection indeed entails a dimensionality reduction process which can be easily illustrated through PCA. We can describe the samples for every individual signal to belong to $\boldsymbol{S}$ and write them as $y_1, y_2, ..., y_{\mathrm{N}} \in \mathbb{F}^{\mathrm{M}}$, where N is the number of patients (or individual signals) in the M-dimensional space. It is a convenient preprocessing step to strip off the mean as $\sum_\kappa y_\kappa = 0$.

The objective is now to maintain the highest amount of variance in the smaller dimensional space, defined by some integer P smaller than M. Thus, we are mapping as $\mathbb{F}^{\mathrm{M}} \longmapsto \mathbb{F}^{\mathrm{P}}$ to represent the VCG in a P-dimensional space. Namely, we are dealing with an optimisation problem of the form:

$$\min_{\tilde{y}} \left(\sum_{\kappa}^{\mathrm{N}} \|y_\kappa\|_2 - \sum_{\kappa}^{\mathrm{N}} \|\tilde{y}_\kappa\|_2 \right)$$

The projection matrix for dimensionality reduction can be defined by Θ such that $\Theta^{\mathrm{H}}\Theta = \mathbb{I}_{\mathrm{P}}$, being then the mapping $\mathbb{F}^{\mathrm{M}} \longrightarrow \mathbb{F}^{\mathrm{P}} : y \longmapsto \tilde{y} = \Theta\Theta^{\mathrm{H}} y$.

2.5. Assessing Closeness in a H.D. Space

Different metrics have been proposed to evaluate the closeness between points, points and distributions and distributions in a high dimensional space [9]. An interesting metric among all those proposed for our problem is the *Mahalanobis Distance* ($\mathbb{d}_{\mathrm{M}}$), which considers the distribution of the samples in the multivariate space, weighting the closeness with the inverse covariance matrix ($Q := \mathrm{K}_{\varsigma\varsigma}^{-1}$) of the cluster – which is naturally non-singular. This is a classic way of dealing with correlating samples in a multivariate space, and can be understood in layman's terms as measuring the relative distance to the centroid of the distribution $\mathscr{J}$, that is, how many standard deviations a point ς (i.e. $\varsigma = (\varsigma_1, \dots, \varsigma_{\mathrm{M}})$) stands from the mean $\boldsymbol{\mu}$ (such that $\boldsymbol{\mu} = (\mu_1, \dots, \mu_{\mathrm{M}})$) of $\mathscr{J}$.

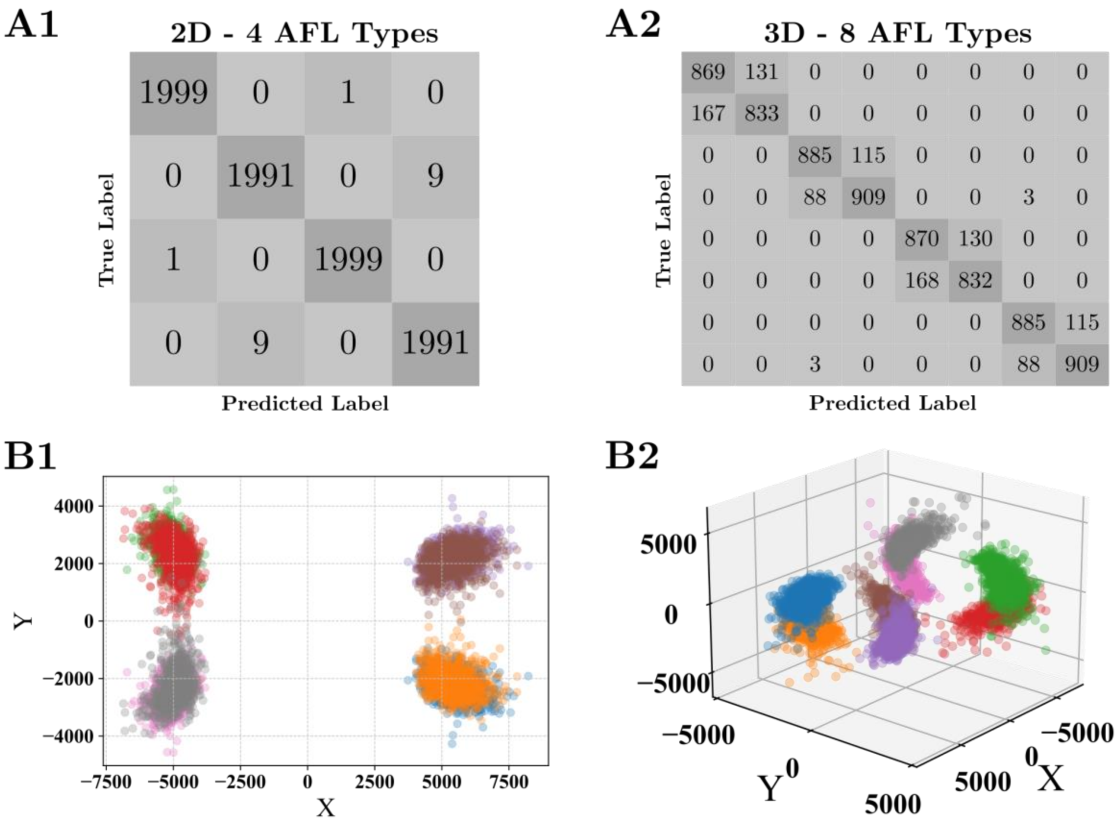

Figure 2. A1-A2: *Confusion matrix of the unsupervised clustering assignment in comparison to the ground truth for 4 (according to slow velocity region) and 8 (slow velocity and conduction direction) AFL types.* **B1-B2:** *Visualisation of a 2D and 3D vector space representation of the VCG signals for each of the 8 groups.*

We can define it as:

$$d_M = \sqrt[2]{(\boldsymbol{\varsigma}-\boldsymbol{\mu})^T Q(\boldsymbol{\varsigma}-\boldsymbol{\mu})}$$

where, for the sake of clarity, the M-dimensional mean vector is defined as:

$$\mu := \mathbb{E}[\boldsymbol{\varsigma}] = (\mathbb{E}[\varsigma_1], \mathbb{E}[\varsigma_1], \dots, \mathbb{E}[\varsigma_M])^T$$

and the precision matrix as:

$$Q := K_{\varsigma_i\varsigma_j}^{-1} = (\mathbb{E}[(\varsigma_i - \mu_i)(\varsigma_j - \mu_j)])^{-1}$$

with $1 \leq i, j \leq M$.

As a final remark, the choice of distance metric over other options such as *Chebyshev* or *Minkowsky* over a normed vector space, was tested on toy problems that evaluated their clustering performance over ground truths based on multivariate Gaussian Distributions with density:

$$f_{\boldsymbol{\varsigma}}(\varsigma_1, \dots, \varsigma_M) = \frac{e^{-\frac{1}{2}(\boldsymbol{\varsigma}-\boldsymbol{\mu})^T Q(\boldsymbol{\varsigma}-\boldsymbol{\mu})}}{\sqrt[2]{(2\pi)^M |K_{\varsigma\varsigma}|}}$$

for the non-degenerate case. The generalisation to the degenerate case comes naturally from defining Q as the Moore-Penrose Pseudoinverse of $K_{\varsigma\varsigma}$, i.e. $Q := K_{\varsigma_i\varsigma_j}^{\dagger}$.

2.6. Hierarchical clustering

To perform a classification, agglomerative nesting techniques are widely used to form groups in a stratified manner. In particular, hierarchical clustering analysis techniques are proposed in this study to group the families of signals that belong to the AFL type.

From all the possible hierarchical clustering algorithms, Ward's version is found to be more efficient when looking for a clear hierarchy or clusters [10]. The implementation relies on the metric:

$$d_W(u,v) = \sqrt{\frac{|v|+|r|}{C} d_W(u,v)^2 + \frac{|v|+|l|}{C} d_W(u,l)^2 - \frac{|v|}{C} d_W(r,l)^2},$$

where the merging of r and l is represented by u and then v is the cluster that is not used for the forest, representing the cardinality with $|\cdot|$ and $C = |v| + |r| + |l|$.

3. Results

3.1. Closeness Among Clusters

The designed metric S is created to define closeness (S=1 infinitely close, S=0 infinitely far). The clusters are evaluated in different low-dimensional sub-spaces projected from the original (579 dimensional-) space. The results of this metric are evaluated over clusters that have same slow conduction region ('velocity') and direction of conduction ('direction'), those with same conduction velocity (ignoring 'direction'), and those with no velocity and conduction direction in common. Results are displayed in Table 1 and illustrations in Figure 2 B1-B2.

Table 1. Distance Metric Results Amongst Groups in Different Reduced-Dimensional Spaces.

Distance Metric Analysis		
Dimensions	Groups	S
	Explained Variance = 0.996	
n = 6	= Velocity, = Direction	0.35
	= Velocity	0.93
	Different Groups	0.04
	Explained Variance = 0.880	
n = 3	= Velocity, = Direction	0.47
	= Velocity	0.94
	Different Groups	0.02
	Explained Variance = 0.820	
n = 2	= Velocity, = Direction	0.94
	= Velocity	0.93
	Different Groups	0.02

3.3. Clustering Analysis

K-means and Gaussian Mixture Models clusterings were used to evaluate the accuracy of the unsupervised classification of AFL types in the vector subspaces. Results for the different metrics are displayed in Table 2 and Figure 2.A1-A2.

Table 2. Clustering Analysis Metrics Results: Macro-precision, recall, F1 score, total true positive, total false positives and total false negatives.

Clustering Analysis		
Dimensions	Parameter	Value
	# Clusters = 8	
n = 3	Macro-Prec = Recall = F1 Sc.	0.874
	Total TP	6998
	Total FP = Total FN	1002
	# Clusters = 4	
n = 3	Macro-Prec = Recall = F1 Sc.	0.999
	Total TP	7994
	Total FP = Total FN	6
	# Clusters = 4	
n = 2	Macro-Prec = Recall = F1 Sc.	0.997
	Total TP	7980
	Total FP = Total FN	20

4. Discussion

Firstly, it is noticeable that principal components from the 7th dimension onwards explained less than 0.01 of the variance, proving the generalisability of the results in a computationally efficient way. Furthermore, the 6D-space results for the S metric show a clearly trained space where new unlabelled samples would be accurately classified, according to its slow-conduction and direction velocity, that is, its AFL type (see tables 1 and 2). This multiclass classification is still effective with 0.88 variance in a 3D space, although in a 2D space the direction component is lost, meaning that we would only classify AFL according to the slow velocity conduction region – e.g. common from perimetral but not common CW from common CCW. In layman's terms, new clinical cases could be diagnosed with high accuracy once this method is applied to train a vector space over previously diagnosed cases.

Finally, the clustering analysis (see Figure 2.A1-A2) illustrates the accuracy of unsupervised classification of the AFL types from the ML techniques applied to the raw VCG signal. In many cases (e.g. AFL), obtaining this information prior to ablation or treatment would be of great use for the physicians.

5. Final Conclusions

This article proves that classic bio-signal analysis techniques combined with ML techniques in dimensional transforms in the context of Hilbert Spaces can provide a novel methodology to non-invasively diagnose and classify cardiac diseases from their ECG. This promising technique, proven to work in previously published simulations, is currently being applied to clinical signals.

Acknowledgments

Thanks to the *Rafael del Pino* foundation, *La Caixa* foundation and *Fung Institute* at UC Berkeley for the contribution to the academic training of the first author of this publication. This work was supported by PID2019-109547RB-I00 (National Research Program, Ministerio de Ciencia e Innovación, Spanish Government), CIBERCV CB16/11/00486 (Instituto de Salud Carlos III) and PROMETEOII/2014/037 (Generalitat Valenciana).

References

[1] A. J. Rogers, A. Selvalingam, M. I. Alhusseini, D. E. Krummen, C. Corrado, F. Abuzaid, T. Baykaner, C. Meyer, P. Clopton and e. a. W. Giles, "Machine learned cellular phenotypes in cardiomyopathy predict sudden death," *Circulation Research*, vol. 128, no. 2, pp. 172-184, 2021.

[2] S. H. Jambukia, V. K. Dabhi and H. B. Prajapati, "Classification of ECG signals using machine learning techniques: A survey," *2015 International Conference on Advances in Computer Engineering and Applications*, pp. 714-721, IEEE, 2015.

[3] B. Deb, P. Ganesan, R. Feng, N. K. Bhatia, A. J. Rogers, S. Ruiperez-Campillo, P. Clopton and S. M. Narayan, "Unsupervised machine learning identifies phenotypes for atrial fibrillation that predict acute ablation success," *Journal of the American College of Cardiology*, vol. 79, no. 9_Supplement, p. 51, 2022.

[4] A. Minchole and B. Rodriguez, "Artificial intelligence for the electrocardiogram," *Nature medicine*, vol. 25, no. 1, pp. 22-23, 2019.

[5] G. Luongo, S. Schuler, A. Luik, T. P. Almeida, D. C. Soriano, O. Dössel and A. Loewe, "Non-invasive characterization of atrial flutter mechanisms using recurrence quantification analysis on the ECG: a computational study," *IEEE Transactions on Biomedical Engineering*, vol. 68, no. 3, pp. 914-925, 2020.

[6] E. Herzog, E. Argulian, S. B. Levy and E. F. Aziz, "Pathway for the management of atrial fibrillation and atrial flutter," *Critical pathways in cardiology*, vol. 16, no. 2, pp. 47-52, 2017.

[7] S. Ruiperez-Campillo, S. Castrejon, M. Martinez, R. Cervigon, J. L. M. O. Meste, J. Millet and F. Castells, "Non-invasive characterisation of macroreentrant atrial tachycardia types from a vectorcardiographic approach with the slow conduction region as a cornerstone," *Computer methods and programs in biomedicine*, vol. 200, no. 3, p. 105932, 2021.

[8] S. Ruiperez-Campillo, S. Castrejon, M. Martinez, R. Cervigon, O. Meste, J. Merino, J. Millet and F. Castells, "Slow conduction regions as a valuable vectorcardiographic parameter for the non-invasive identification of atrial flutter types," *2020 Computing in Cardiology*, pp. 1-4, IEEE, 2020.

[9] C. Donnat and H. Susan, "Tracking network dynamics: A survey of distances and similarity metrics.," *arXiv preprint:1801.07351*, 22 January 2018.

[10] S. C. Johnson, "Hierarchical clustering schemes," *Psychometrika* , vol. 32, no. 3, pp. 241-254, 1967.

Address of Correspondence

Samuel Ruipérez-Campillo – School of Medicine and
Engineering, Stanford University, 94305 Palo Alto, CA, USA.
sruiperez@berkeley.edu