Catheter ablation of persistent atrial fibrillation guided by electrograms spatiotemporal dispersion : automatic identification using machine learning approaches

Amina Ghrissi Abdelkader

UNIVERSITÉ CÔTE D'AZUR

ÉCOLE DOCTORALE
SCIENCES ET TECHNOLOGIES DE L'INFORMATION ET DE LA COMMUNICATION

THÈSE DE DOCTORAT

Ablation par Catheter de Fibrillation Atriale Persistante Guidée par Dispersion Spatiotemporelle d'Electrogrammes : Identification Automatique Basée sur l'Apprentissage Statistique

Amina Ghrissi

Laboratoire d'Informatique, Signaux et Systèmes de Sophia Antipolis (I3S)
UMR7271 Université Côte d'Azur CNRS

Présentée en vue de l'obtention du grade de docteur en Automatique, Traitement du Signal et des Images **de l'**Université Côte d'Azur
Dirigée par:
Vicente Zarzoso, Johan Montagnat
Soutenue le: 19 Avril 2021

Devant le jury, composé de:
Vicente Zarzoso, Directeur de thèse, Université Côte d'Azur
Johan Montagnat, Co-Directeur de thèse, Université Côte d'Azur
Saeid Sanei, Rapporteur, U. de Notthigham Trent, Royaume-Uni
Rémi Dubois, Rapporteur, IHU LIRYC
Flavia Ravelli, Examinatrice, Université de Trento, Italie
Bertrand Rivet, Examinateur, Université Grenoble Alpes
Charles Bouveyron, Président, Université Côte d'Azur

THÈSE DE DOCTORAT

UNIVERSITÉ CÔTE D'AZUR

ÉCOLE DOCTORALE STIC

Sciences et Technologies de l'Information et de la Communication

Laboratoire I3S, CNRS

Ablation par Catheter de Fibrillation Atriale Persistante Guidée par Dispersion Spatiotemporelle d'Electrogrammes : Identification Automatique Basée sur l'Apprentissage Statistique

Thèse pour obtenir le grade de

Docteur en Automatique, Traitement du Signal et des Images

Présentée par
Amina Ghrissi

Dirigée par
Vicente Zarzoso, Professeur, I3S, CNRS, Université Côte d'Azur
Johan Montagnat, Directeur de Recherche, I3S, CNRS, Université Côte d'Azur

Soutenue publiquement le 19 Avril 2021 devant le jury composé de:

Président:	Charles Bouveyron	Professeur, Inria, Université Côte d'Azur
Rapporteur:	Saeid Sanei	Professeur, University de Notthigham Trent, Royaume-Uni
Rapporteur:	Rémi Dubois	Professeur, IHU LIRYC
Examinatrice:	Flavia Ravelli	Professeure, Université de Trento, Italie
Examinateur:	Bertrand Rivet	Maître de Conférences, Université Grenoble Alpes
Directeur de thèse:	Vicente Zarzoso	Professeur, I3S, CNRS, Université Côte d'Azur
Co-Directeur de thèse:	Johan Montagnat	Directeur de Recherche, I3S, CNRS, Université Côte d'Azur

Résumé

La fibrillation atriale (FA) est l'arythmie cardiaque soutenue la plus fréquemment rencontrée dans la pratique clinique. Pour la traiter, l'ablation par cathéter de zones cardiaques jugées responsables de soutenir l'arythmie est devenue la thérapie la plus utilisée. Un nouveau protocole d'ablation se base sur l'identification des zones atriales où les électrogrammes (EGM) enregistrés à l'aide d'un cathéter à électrodes multiples, appelé PentaRay, manifestent des décalages spatiotemporels significatifs sur plusieurs voies adjacentes. Ce phénomène est appelé dispersion spatio-temporelle (DST). L'intervention devient ainsi plus adaptée aux spécificités de chaque patient et elle atteint un taux de succès procédural de 95%. Cependant, à l'heure actuelle les zones de DST sont identifiées de manière visuelle par le spécialiste pratiquant l'ablation. Cette thèse vise à identifier automatiquement les sites potentiels d'ablation basée sur la DST à l'aide de techniques d'apprentissage statistique et notamment d'apprentissage profond adaptées. Dans la première partie, les enregistrements EGM sont classés par catégorie en DST vs. non-DST. Cependant, le rapport très déséquilibré entre les données issues des deux classes dégrade les résultats de classification. Nous abordons ce problème en utilisant des techniques d'augmentation de données adaptées à la problématique médicale et qui permettent d'obtenir de bons taux de classification. La performance globale s'élève ainsi atteignant des valeurs de précision et d'aire sous la courbe ROC autour de 90%. Deux approches sont ensuite comparées, l'ingénierie des caractéristiques et l'extraction automatique de ces caractéristiques par apprentissage statistique à partir d'une série temporelle, appelée valeur absolue de tension maximale aux branches du PentRay (VAVp). Les résultats montrent que la classification supervisée de VAVp est prometteuse avec des valeurs de précision, sensibilité et spécificité autour de 90%. Ensuite, la classification des enregistrements EGM bruts est effectuée à l'aide de plusieurs outils d'apprentissage statistique. Une première approche consiste à étudier les circuits arithmétiques à convolution pour leur intérêt théorique prometteur, mais les expériences sur des données synthétiques sont infructueuses. Enfin, nous investiguons des outils d'apprentissage supervisé plus conventionnels comme les réseaux de neurones convolutifs (RNC). Nous concevons une sélection de représentation des données adaptées à différents algorithmes

de classification. Ces modèles sont ensuite évalués en termes de performance et coût de calcul. L'apprentissage profond par transfert est aussi étudié. La meilleure performance est obtenue avec un RNC peu profond pour la classification des matrices EGM brutes, atteignant 94% de précision et d'aire sous la courbe ROC en plus d'un score F1 de 60%. Dans la deuxième partie, les enregistrements EGM acquis pendant la cartographie sont étiquetés ablatés vs. non-ablatés en fonction de leur proximité par rapport aux sites d'ablation, puis classés dans les mêmes catégories. Les annotations de dispersion sont aussi prises en compte comme une probabilité à priori dans la classification. La meilleure performance représente un score F1 de 76%. L'agrégation de l'étiquette DST ne permet pas d'améliorer les performances du modèle. Globalement, ce travail fait partie des premières tentatives d'application de l'analyse statistique et d'outils d'apprentissage pour l'identification automatique et réussie des zones d'ablation en se basant sur la DST. En fournissant aux cardiologues interventionnels un outil intelligent, objectif et déployé en temps réel qui permet la caractérisation de la dispersion spatiotemporelle, notre solution permet d'améliorer potentiellement l'efficacité de la thérapie personnalisée d'ablation par cathéter de la FA persistante.

Mots clés: fibrillation atriale persistante, dispersion spatiotemporelle, ablation, cathéter multipolaire PentaRay, électrogramme multicanal, valeur absolue de tension maximale, apprentissage statistique, classification, partitionnement de données, augmentation des données, apprentissage par transfert, réseau de neurones convolutifs.

Abstract

Catheter ablation is increasingly used to treat atrial fibrillation (AF), the most common sustained cardiac arrhythmia encountered in clinical practice. A recent patient-tailored AF ablation therapy, giving 95% of procedural success rate, is based on the use of a multipolar mapping catheter called PentaRay. It targets areas of spatiotemporal dispersion (STD) in the atria as potential AF drivers. STD stands for a delay of the cardiac activation observed in intracardiac electrograms (EGMs) across contiguous leads. In practice, interventional cardiologists localize STD sites visually using the PentaRay multipolar mapping catheter. This thesis aims to automatically characterize and identify ablation sites in STD-based ablation of persistent AF using machine learning (ML) including deep learning (DL) techniques. In the first part, EGM recordings are classified into STD vs. non-STD groups. However, highly imbalanced dataset ratio hampers the classification performance. We tackle this issue by using adapted data augmentation techniques that help achieve good classification. The overall performance is high with values of accuracy and AUC around 90%. First, two approaches are benchmarked, feature engineering and automatic feature extraction from a time series, called maximal voltage absolute values at any of the bipoles (VAVp). Statistical features are extracted and fed to ML classifiers but no important dissimilarity is obtained between STD and non-STD categories. Results show that the supervised classification of raw VAVp time series itself into the same categories is promising with values of accuracy, AUC, sensitivity and specificity around 90%. Second, the classification of raw multichannel EGM recordings is performed. Shallow convolutional arithmetic circuits are investigated for their promising theoretical interest but experimental results on synthetic data are unsuccessful. Then, we move forward to more conventional supervised ML tools. We design a selection of data representations adapted to different ML and DL models, and benchmark their performance in terms of classification and computational cost. Transfer learning is also assessed. The best performance is achieved with a convolutional neural network (CNN) model for classifying raw EGM matrices. The average performance over cross-validation reaches 94% of accuracy and AUC added to an F1-score of 60%. In the second part, EGM recordings acquired during mapping are labeled ablated

vs. non-ablated according to their proximity to the ablation sites then classified into the same categories. STD labels, previously defined by interventional cardiologists at the ablation procedure, are also aggregated as a prior probability in the classification task. Classification results on the test set show that a shallow CNN gives the best performance with an F1-score of 76%. Aggregating STD label does not help improve the model's performance. Overall, this work is among the first attempts at the application of statistical analysis and ML tools to automatically identify successful ablation areas in STD-based ablation. By providing interventional cardiologists with a real-time objective measure of STD, the proposed solution offers the potential to improve the efficiency and effectiveness of this fully patient-tailored catheter ablation approach for treating persistent AF.

Keywords: persistent atrial fibrillation, spatiotemporal dispersion, ablation, PentaRay multipolar catheter, multichannel electrogram, maximal voltage absolute values, machine learning, classification, clustering, data augmentation, transfer learning, convolutional neural network.

I dedicate this work to my mother Latifa

Acknowledgements

My deepest greetings go to all the people who helped me in the elaboration of this doctoral thesis. I would like to express my sincere thanks to the members of the jury for accepting to evaluate and examine this work and for the reviewers for being patient with my thesis manuscript.

I offer a lot of gratitude and appreciation to Université Côte d'Azur for funding my reaserch for more than three years and to I3S laboratory that held my work and afforded me with technical support, supervision and team integration. I would particularly like to single out my thesis director Prof. Vicente Zarzoso. I have been amazingly fortunate to have an advisor who gave me the freedom to explore on my own, and at the same time the guidance to recover when my steps faltered. He taught me how to question thoughts and express ideas. Gracias Vicente! I would like also to thank my thesis co-supervisor, DR. Johan Montagnat who afforded me with guidance and help at the difficult moments. Special thanks go to Prof. Olivier Meste for his insightful feedback and support, especially when a contract extension was required. Special thanks go to IADB project coordinators, mainly Prof. Frédéric Précioso who afforded me with expertise on machine learning while being particularly kind.

My deepest gratitude and acknowledgment go to our partner interventional cardiologist, Dr. Fabien Squara from Nice CHU, without whom this work wouldn't have been possible. Fabien proposed the topic of the identification of ablation sites in STD-based ablation. He also afforded us with real EGM data. Thanks to the feedback of Fabien, we were able to continuously enhance our proposed solution and to adapt it to the medical topic. Again thanks to Fabien, the collaboration with technicians from Biosense Webster company was possible. My greetings go to Bilel Rezig, Vincent Roger for helping us in understanding the structure of the exported data.

I would like to present my deepest thanks and appreciation to Dr. Douglas Almonfrey for his excellent cooperation and for his invaluable remarks and advice. His tutorship offered a critical assessment of some solutions I developed and the way to implement them.

I would like to mention that one of the most exiting events in my PhD, was the day

Dr. Gabriel Lactu, cardiologist at Centre Hospitalier Princesse Grace (CHPG) gave me the opportunity to attend an AF ablation intervention. The discussions with the CHPG team, composed of doctors and researchers, gave me an objective feedback about my work and taught me more about the challenges of AF ablation.

Special thanks go also to my office colleagues Pedro Marinho Ramos de Oliveira and previously Jean-Marie Kai for the funny moments, the delicious cups of tea and the mad trips we shared together. I would never forget the invaluable help of Ninad Manerikar with the tricky installation of CUDA. It was a nightmare without your help!

To the dream team, as the project manager of Invent@UCA used to call us, thank you all for the funny moments we shared together and for being cooperative and productive even under the pressure of competition. We do deserve that prize.

From Tunisia, I am grateful to all my teachers and professors who stimulated my curiosity and my devotion to science. Participating to a data science challenge and being selected in the middle of the PhD battle was a dream but very hard to accomplish. Hopefully, Marwen Sallem made it come true.

Finally and most importantly, none of this would have been possible without the love and infinite support of my dear parents, cutest brother on the earth and all my family members, in my native country, to whom this work is dedicated. No words can describe the devotion of my mother Latifa Belhadj Ali to my education and empowerment. Only you mom can understand this *"Bentoutou"*.

And to make sure the PhD goes well at home, I recommend getting married with a lovely professor. My deepest greetings and appreciation go to my husband Dr. Dhafer Abdelkader, who believed in my potential, made sacrifices to help me succeed and afforded me with unlimited love and affection. Sorry for spending half of the honeymoon working on my PhD and for spending the weekends exporting data from the hospital, but as you always say *"Hard work pays off"*.

Catheter Ablation of Persistent Atrial Fibrillation Guided by Electrograms Spatiotemporal Dispersion: Automatic Identification Using Machine Learning Approaches

Amina Ghrissi

Notation

We denote scalars, vectors, matrices and tensors respectively by lower-case (a,b,c,...), boldface lower-case (**a**,**b**,**c**,...), boldface upper-case (**A**,**B**,**C**,...) and calligraphic upper-case ($\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$,...) letters.

$\|.\|$	–	absolute value
$\|\|.\|\|_2$	–	ℓ_2 Euclidean norm
$E(.)$	–	mathematical expectation
$[M]$	–	$\{1, 2, \ldots, M\}$
diag$(\cdot)$	–	builds a diagonal matrix by placing elements of the input vector along the diagonal
L^2	–	Lebesgue function space
.	–	scalar product $\langle \cdot, \cdot \rangle$
$\otimes$	–	outer product
rank(.)	–	tensor rank
$\mathbb{Q}$	–	Set of rational numbers
$\mathcal{N}(\mu, \sigma)$	–	Normal distribution of mean μ and standard deviation σ

Acronyms

1D	–	one-dimensional
2D	–	two-dimensional
3D	–	three-dimensional
AA	–	Atrial Activity
AF	–	Atrial Fibrillation
AI	–	Artificial Intelligence
AP	–	Action Potential
AS	–	Apprentissage Statistique
AT	–	Atrial Tachycardia
AV	–	Atrioventricular node
BP	–	Backpropagation
bpm	–	Beats Per Minute
CA	–	Catheter Ablation
CHU	–	University Hospital Center
CNN	–	Convolutional Neural Network
CIR	–	Class Imbalance Ratio
Conv	–	Convolutional
ConvAC	–	Convolutional Arithmetic Circuit
CPD	–	Canonical Polyadic Decomposition
DL	–	Deep Learning
DNN	–	Deep Neural Network
EEC	–	External Electrical Cardioversion
ECG	–	Electrocardiogram
EP	–	Electrophysiology

EGM	–	Electrogram
FC	–	Fully Connected
GB	–	Gigabytes
GD	–	Gradient Descent
GAN	–	Generative Adversarial Network
GPU	–	Graphics Processing Unit
HCA	–	Hierarchical Clustering Analysis
LA	–	Left Atrium
LDA	–	Linear Discriminant Analysis
LMS	–	Least Mean Squares
MC	–	Monte Carlo
ML	–	Machine learning
MLR	–	Multivariate Logistic Regression
NMSE	–	Normalized Mean Squared Error
NSR	–	Normal Sinus Rhythm
PCA	–	Principal Component Analysis
PDF	–	Probability Density Function
PV	–	Pulmonary Vein
PVI	–	Pulmonary Vein Isolation
RA	–	Right Atrium
ReLU	–	Rectified Linear Unit
SA	–	Sinoatrial node
SGD	–	Stochastic Gradient Descent
STD	–	Spatiotemporal Dispersion
stdev	–	standard deviation
SVM	–	Support Vector Machine
VA	–	Ventricular Activity

Contents

List of Figures

List of Tables

Part I

Introduction

Chapter 1

Thesis Overview

Contents

1.1 Clinical Context

Atrial Fibrillation (AF) represents the most frequent sustained arrhythmia experienced in clinical practice, rising in prevalence with advancing age. It is considered as the last great frontier of cardiac electrophysiology as it continues to puzzle cardiologists [1]. According to [2], the spreading rhythm of AF can be assimilated to a new epidemic and reveals a public health challenge. This complex arrhythmia is associated with a five-fold increase in stroke risk [3] [4] [5]. Among the existing treatments of persistent AF, ablation interventions prove more efficient than drug therapies in terms of long-term success. Ablation consists in burning cardiac tissue areas thought to be responsible for the genesis and maintenance of AF [6]. Ablation catheters are used to

deliver radiofrequency (RF) energy at these specific atrial sites. The classical ablation approach is called stepwise but it yields poor clinical results [7] [8]. A novel wholly patient-tailored ablation protocol, giving 95% of procedural success rate, consists in identifying ablation sites based on a pattern called spatiotemporal dispersion (STD). Multipolar mapping catheters are used to record electrograms (EGM) in the atria thus targeting areas of STD as potential AF drivers. The high-density mapping PentaRay catheter is used for STD localization. It has a five-branch star design with two bipoles on each spline. According to preliminary guidelines for STD identification from visual inspection, the 10-channel EGM recording acquired by the PentaRay would display a cardiac activation delay of 70% of AF cycle length on a minimum of three neighboring bipoles (channels) [9]. However, this visual inspection may be biased by the difficulty of the interventional cardiologist to quantify the STD pattern at each single mapped location in real time and accurately, as thousands of cardiac sites are checked in a typical ablation intervention.

1.2 Objectives

The present PhD thesis aims to design a decision-aid solution based on artificial intelligence [10] that would help interventional cardiologists to automatically identify potential target sites in STD-based ablation therapy of persistent AF.

1.2.1 General Objectives

The main contribution of the present doctoral thesis is three-fold:

1. Building a structured dataset of EGM recordings of interest.
2. Designing a solution that identifies automatically STD EGMs while handling the highly imbalanced dataset issue.
3. Detecting potential ablation sites in the atria automatically.

To our knowledge, this contribution applies for the first time automatic feature selection and deep learning (DL) for the automatic identification of dispersion pattern in multi-channel EGMs and the detection of potential ablation sites.

1.2.2 Specific Objectives

The specific goals are detailed as follows:

- Handling the highly imbalanced dataset issue between STD and non-STD samples with adapted data augmentation (DA) methods.

- Classifying multichannel EGM recordings into STD vs. non-STD categories.
- Investigating the use of convolutional arithmetic circuits, a promising type of shallow convolutional neural network (CNN), whose transfer function can be formulated as a tensor decomposition, in the problem of STD detection.
- Investigating statistical and handcrafted features for STD identification.
- Benchmarking several ML models and settings for automatic STD detection in order to reach the best performance while ensuring low inference time, thus allowing on-the-fly detection of dispersion pattern. STD identification passes through investigating several feature selection methods adapted to each ML architecture.
- Applying transfer learning with the use of a deep CNN and a specific data transformation as a proof of concept.
- Designing an annotation model to identify mapped locations that are ablated.
- Classifying multichannel EGM recordings into ablated vs. non-ablated while taking into account both STD prior probability and some physiological information.

1.3 Study Datasets

The University Hospital Center (CHU) of Nice affords the study with real data acquired from ablation interventions of persistent AF patients. All patients are ablated with STD-based protocol. Data are annotated by interventional cardiologists as "STD" or others. The raw data acquired should be restructured, preprocessed and revised before experiments as explained in Sec. 5.2.

Added to real signals, we design a realistic EGM model that includes STD pattern with different settings. Synthetic signals are then generated for preliminary experiments in STD detection.

1.4 Manuscript Organization

The remaining of the manuscript consists of 9 chapters (Chapters 2 – 10) grouped in 3 parts. Part I includes the present chapter and other 3 chapters (Chapters 2 – 4) and presents a theoretical background necessary for the comprehension of this thesis, as well as the fundamental state-of-the-art that supports and motivates the contributions reported later. Part II consists of 5 chapters (Chapters 5 – 9) and presents the original contributions of the present doctoral thesis. Finally, Part III concludes this work and consists of a single chapter, that is, Chapter 10. More specifically:

- **Chapter 2** explains the cardiac anatomy, electrophysiology and conduction system of the human heart. Particular attention is given to presenting AF as the clinically most common sustained arrhythmia, first by presenting statistics about AF then explaining its underlying mechanisms and potential therapies like catheter ablation. EGM is also defined as a recording of intracardiac electrical activity.
- **Chapter 3** draws the state of art about catheter ablation of AF with RF energy and introduces STD-based ablation. A description of the PentaRay catheter is also provided.
- **Chapter 4** reviews the mathematical background needed to understand the statistical analysis and ML experiments conducted in the present doctoral thesis. The chapter starts by recalling the basics of classification algorithms and performance metrics. Deeper architectures like VGG16 convolutional neural network are then presented.
- **Chapter 5** describes the study datasets starting with the real set. The process of data exportation, anonymization and structuring is detailed in addition to statistics about the study population. Then, a realistic multichannel EGM signal model is explained while showing some synthetic 10-channel signals with different settings of fractionation degrees and delay values.
- **Chapter 6** presents one of the first attempts to classify the multichannel EGM recordings into STD vs. non-STD categories with the use of baseline ML algorithms. The main challenges encountered, in the classification task, consists in the high class imbalance ratio that will be addressed with adapted DA methods.
- **Chapter 7** benchmarks two approaches for the characterization and identification of STD EGMs: *handcrafted feature extraction* and *automatic feature extraction*. A statistical analysis, including the calculation of kurtosis and data distribution, is first performed on synthetic STD EGMs then to the real one. A compact representation of data is used in this chapter.
- **Chapter 8** represents a comparative study of different classification models of the EGM recordings into STD vs. non-STD. Each model consists in a way of representing the multichannel EGM data combined with an adapted ML algorithm. The algorithms include baseline classifiers, end-to-end trained CNNs added to transfer learning applied to a deep CNN architecture.
- **Chapter 9** proposes to automatically identify potential ablation sites in the atria in STD-based ablation. For this purpose, EGM recordings acquired during the mapping phase of the ablation intervention are automatically classified into ablated

vs. non-ablated with the use of ML algorithms. The STD label is also taken into consideration as an additional input in the classification task.

- **Chapter 10** concludes the manuscript by summarizing the study that has been conducted and discussing the experimental results that validate the presented contributions. Then, the present doctoral thesis is ended by focusing on potential new perspectives of automatizing the identification of ablation sites in STD-based ablation of persistent AF analysis.

1.5 Funding & Collaborations

This doctoral thesis is part of the project part of the project Integration and Analysis of Biomedical Data (IADB), partly funded by the French government labelled PIA program under its IDEX UCA^{JEDI} project (ANR-15-IDEX-0001). The PhD is supported by the "Bourse d'excellence" UCA and benefits from an extension of contract co-funded by the I3S laboratory. This work is the fruit of a joint collaboration between Laboratoire d'Informatique, Signaux et Systèmes de Sophia Antipolis (I3S), the cardiology department of CHU of Nice (Pasteur I), Biosense Webster Inc (Irvine, CA, USA) and Instituto Federal do Espírito Santo (IFES, Brazil). The real data are acquired from CARTO system installed at CHU of Nice thanks to our partner cardiologist Dr. Fabien Squara who afforded us with medical expertise and continuous feedback about adapting ML techniques to the physiological constraints of AF. Technicians and engineers from Biosense provided us with technical support for the exportation, exploitation and interpretation of the exported data. Our partner from IFES, Dr. Douglas Almonfrey, afforded technical help about ML algorithms and implementation. He also allowed us to run ML experiments on a powerful GPU (Quadro P6000) at IFES.

1.6 Scientific Productions

1.6.1 Declaration of Invention

Our declaration of invention is currently under examination. Indeed, we have been contacted by a medtech startup who is interested in our work.

1.6.2 Publications

International Journal

- A. Ghrissi, D. Almonfrey, F. Squara, V. Zarzoso and J. Montagnat. "Automatic Detection of Spatiotemporal Dispersion Electrograms in Atrial Fibrillation Ablation

Using Machine Learning: A comparative Study," submitted, Jan. 2021.

International Conferences

- A. Ghrissi, D. Almonfrey, R. Almeida, F. Squara, V. Zarzoso and J. Montagnat, "Data Augmentation for Automatic Identification of Spatiotemporal Dispersion Electrograms in Atrial Fibrillation Ablation Using Machine Learning," in: *Proc. 42th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2020*, Montreal, CA 2020.
  Paper selected as Europe geographic finalist in the student paper contest.

- A. Ghrissi, F. Squara, V. Zarzoso and J. Montagnat, "Identification of Spatiotemporal Dispersion Electrograms in Persistent Atrial Fibrillation Ablation Using Maximal Voltage Absolute Values," in: *Proc. 28th European Signal Processing Conference*, Amsterdam, Netherlands, 2021.

- A. Ghrissi, F. Squara, V. Zarzoso and J. Montagnat, "Identification of Ablation Sites in Persistent Atrial Fibrillation Based on Spatiotemporal Dispersion of Electrograms Using Machine Learning," in: *Proc. Computing in Cardiology Conference*, Rimini, Italy, 2021.

- M. Sallem, A. Ghrissi, A. Saadaoui and V. Zarzoso, "Detection of Cardiac Arrhythmias from Varied Length Multichannel Electrocardiogram Recordings Using Deep Convolutional Neural Networks," PhysioNet/CinC Challenge, in: *Proc. Computing in Cardiology Conference*, Rimini, Italy, 2020.

1.6.3 Other Materials Derived From This Work

Added to our publications, we presented two posters and participated in two challenges that led to a publication and a transfer of intellectual property.

Posters

- A. Ghrissi, F. Squara, J. Montagnat and V. Zarzoso, "Atrial Fibrillation Ablation Guided by Electrogram Spatiotemporal Dispersion using Machine Learning," in: *Atrial Signals, Physicians meet Engineers, poster presentation*, France, 2019.

- A. Ghrissi, J. Montagnat and V. Zarzoso, "Training Neural Networks: A Tensor Based Approach," in: *EURASIP Summer School, Tensor-Based Signal Processing*, poster presentation, Belgium, 2018.

Challenges

- PhysioNet/CinC Challenge, Computing in Cardiology, 2020.
- Entrepreneurial skills are developed thanks to INVENT@UCA, a program that groups professionals, academics and students to support the digital transformation of organizations while training students in innovation and entrepreneurship. The solution developed consists in a mathematical model of the effect of adapted physical activity on elderly people with respect to socio-economical key performance indicators. The designed solution earned the Best Pitch award and the intellectual property was transfered to ORPEA group.

Chapter 2

Cardiac Electrophysiology

Contents

2.1 Introduction

The previous chapter gives an overview of the medical context of the thesis and enumerates the goals and the main contributions achieved. The human heart anatomy and physiology are reviewed in the current chapter. The heart is a muscle that pumps the blood to the different organs of the body and we are interested in the electromagnetic phenomena associated with this vital muscle (Sec: 2.1). The intracardiac electrical ac-

tivity is recorded through mapping catheters and is referred to as electrogram (EGM). Sec. 2.5 is then devoted to AF, the most common cardiac arrhythmia whose mechanisms are still unclear.

2.2 Anatomy and Physiology of the Heart

2.2.1 Basic Structure

The heart is a muscular organ located in the chest between the lungs and functions as the circulatory pump of the body. It is about the size of a closed fist, weighting approximatively 250 to 300 gr. Figure 2.1 gives an overall view of the heart different compartments and surrounding great vessels. The heart is composed of four main chambers:

1. Two upper chambers called atria.
2. Two lower chambers called ventricles.

The blood circulation is ensured through autonomous gates separating the chambers and principal veins, called valves. When the heart beats, or contracts, four valves open and close to let blood flow in only one direction. Each valve is characterized by its specific flaps. The difference in pressure among and across the valves ensures their opening and closing in a highly organized rhythm. The four valves are as follow:

1. Tricuspid valve, located between the right atrium and right ventricle.
2. Pulmonary valve, between the right ventricle and the pulmonary artery.
3. Mitral valve, between the left atrium and left ventricle.
4. Aortic valve, between the left ventricle and the aorta.

Regarding the pumping function, this natural circulatory pump receives in the left atria (LA) the oxygenated blood from the lungs through the pulmonary veins. Then, blood travels through the mitral valve to the left ventricle and to the systemic circulation. This process consists in pumping treated blood and irrigating the body, thus providing oxygen and nutrients to body tissues [11].

2.2.2 Conduction System of the Heart

The heart consists also of two nodes responsible for the conduction system:

1. Sinoatrial node (SA node):
   Located in the right atrium at the superior vena cava, SA consists of specialized muscle cells. The SA nodal cells are self-excitatory and referred to as the natural pacemaker of the heart.

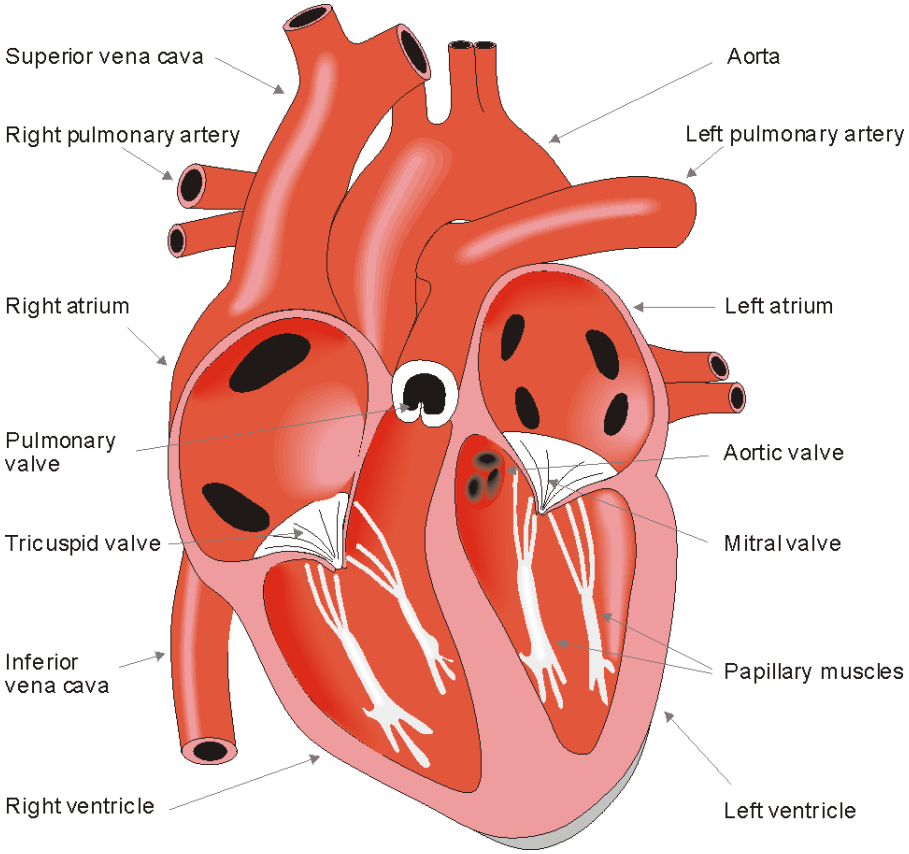

Figure 2.1: The anatomy of the heart and associated vessels. Image taken from [11].

2. Atrioventricular node (AV node):
   Located at the boundary between the atria and ventricles, AV has an intrinsic frequency of about 50 pulses/min that can adapt to higher frequency if it is triggered with a higher pulse.

The AV and SA nodes play a crucial role in the conduction system of the heart. Indeed the natural pacemaker (SA node) emits an electrical excitation to initiate a heart beat. This excitation translates into a propagating depolarization wavefront from the top to the bottom of the atria, a contraction of the atria and flow of blood towards the ventricles. The AV node enters the process and delays the excitation propagation to maintain an efficient beating rhythm. The electrical excitation travels the atria, bundle, Purkinje fibers then reaches the ventricular myocardium. Finally the excitation propagates from the apex to the base of the ventricles making them contract and eject blood. These steps induce the heart beat genesis. In a healthy physiological state, the process of beating is known as normal sinus rhythm (NSR). This process can be decomposed of a sequence of three phases: 1) creation of potential difference at the scale of myocytes; 2) electrical excitation propagation from cell to cell; 3) the aggregation of all electrical activities to generate the final exciting impulse [12].

1. At the scale of myocytes (the heart muscle cells): The electric activation is due to a difference in potential between the intra and extra physiological (chemical) environments as a consequence of the inflow/outflow of potassium (K^+), calcium (Ca^{2+}), chloride (Cl^-) and Sodium (Na^+) ions across the cell membrane through ion channels. Figure 2.2 displays the estimated membrane action potential (AP) of about 100 mV. It presents two main phases

   (a) A plateau phase follows cardiac depolarization.

   (b) Repolarization is a consequence of the outflow of K^+ ions.

2. Cell-to-cell activation: The SA node is a natural pacemaker as its nodal cells are self-excitatory. They generate an AP at the rate of about 70 pulses/min. Thus, the activation propagates throughout the atria following a specific conducting system. Under normal heart beating conditions, the waveform needs to pass through regulator path, mainly the AV node. Normally, the AV node should be excited only by pulses that propagate through to the ventricles. The conduction system is also composed of a common bundle then it separates into two bundle branches (right and left). More distally, the bundles ramify into Purkinje fibers that vehicle the exciting signal to the inner sides of the ventricles.

3. Genesis of electrical activity of the heart out of different wavefronts:

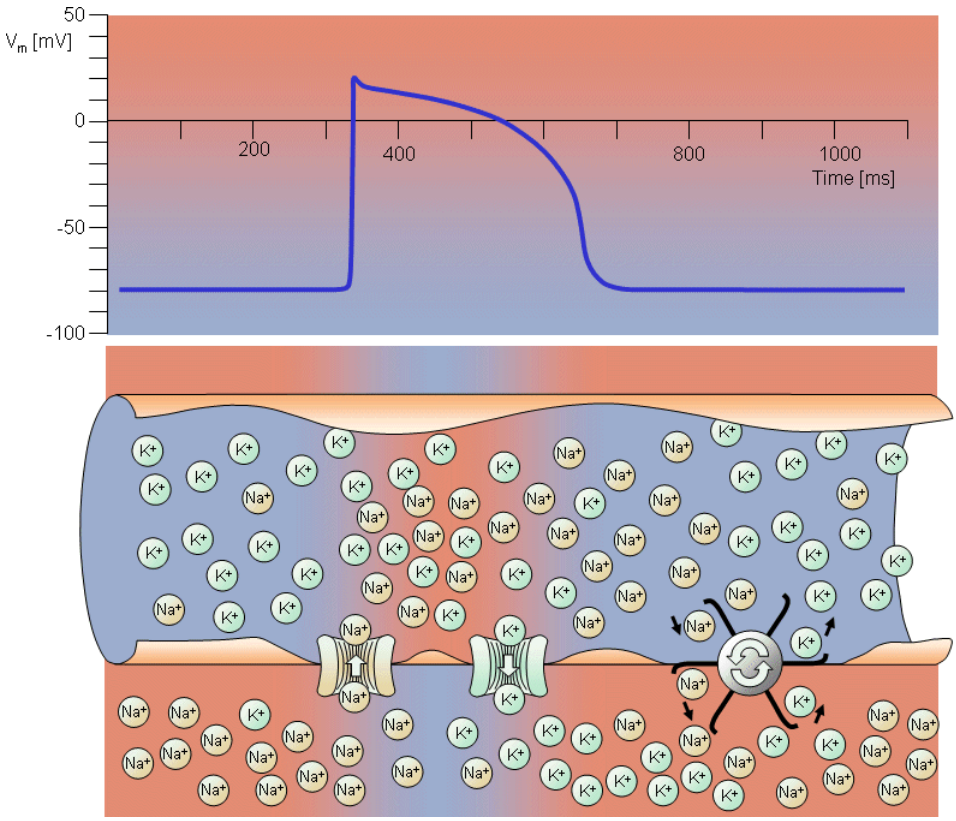

Figure 2.2: Electrophysiology of the cardiac muscle cell. Genesis of an AP as a result of inflow/outflow of ions through gap junctions. Image taken from [11].

From the latter point, the set of activated sites cause the formation of a wavefront that travels to the outer wall throughout the ventricles. After each ventricular muscle region has depolarized, repolarization occurs. Finally, the measured overall heart activity is the summation of all the generated waveforms due to cell-to-cell activation in both depolarization and repolarization phases. This phenomena is detailed in Fig. 2.3.

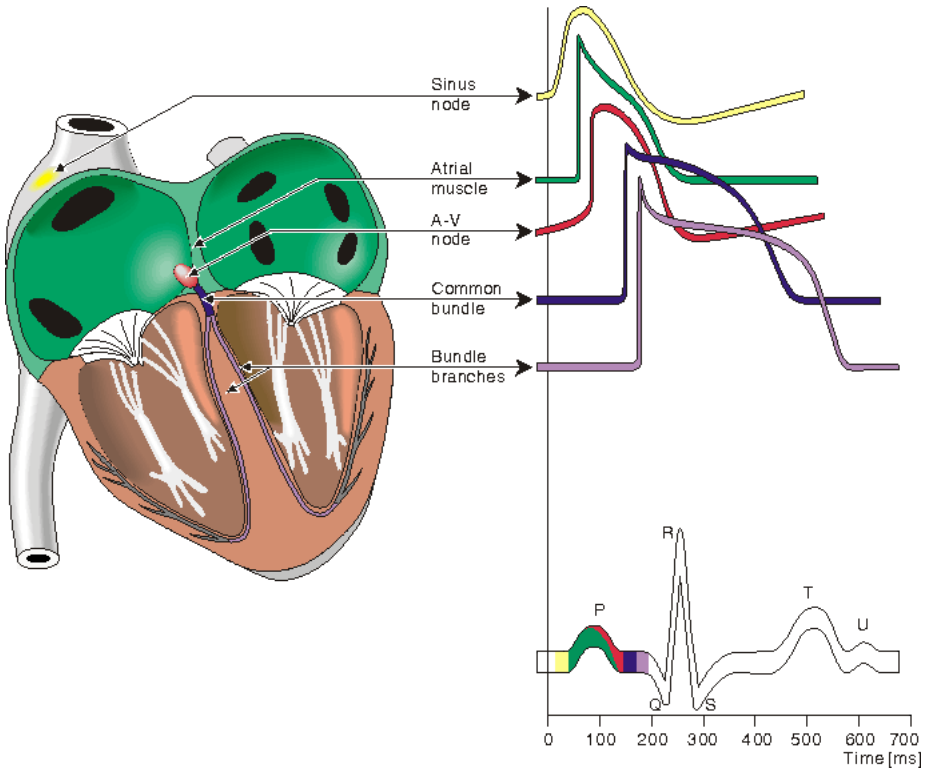

Figure 2.3: The genesis of electric heart activity out of different waveforms for each of the specialized myocytes. Image taken from [11].

2.3 Cardiac Electrical Activity Recording

2.3.1 Electrocardiograms

Contrarily to invasive techniques of heart diagnosis, the ECG records the cardiac electric potentials observable from the body surface, as measured in standard locations on the skin [1]. Based on the difference in potential measured on the different configuration leads and considering the description of overall heart electrical activity detailed in Section 2.2.2, the ECG displays the PQRST waveform shown in Fig. 2.4. The ECG is composed of various segments and intervals, denoted PQRST in alphabetic order as follows [11]:

1. P-wave: atrial depolarization.
2. QRS complex: ventricular depolarization.
3. T-wave: ventricular repolarization.

Atrial repolarization occurs during the ventricular depolarization too (QRS complex) but it is typically masked in normal ECG due to its low amplitude.

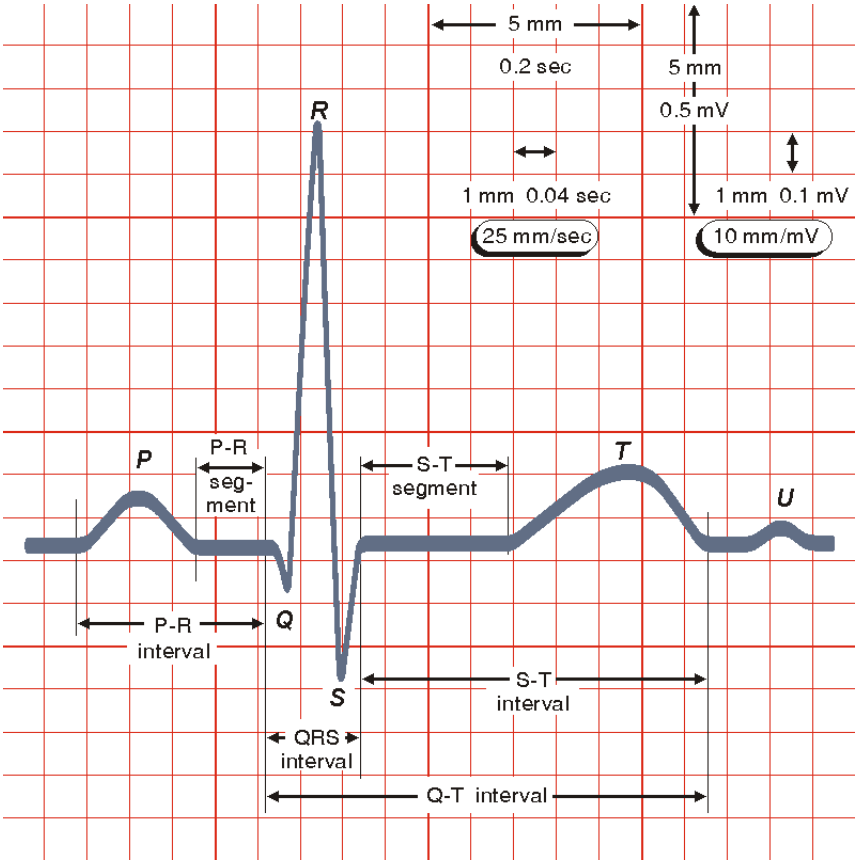

Figure 2.4: The normal ECG. Image taken from [11].

2.3.2 Electrograms

An EGM refers to the signal of the intracardiac electrical local activity of the heart recorded by an endocavity probe during an electrophysiological exploration (mapping). The understanding and interpretation of EGM recordings is fundamental in any electrophysiological investigation and treatment of cardiac arrhythmia. EGMs are generated by a voltage (potential) difference between two electrodes during the cardiac cycle. It is calculated as the differential signal between one electrode connected to the (positive) an-

ode of the amplifier and a second electrode connected to a cathode (negative) input [13]. There exists two types of EGM recordings:

- **Unipolar**
  A unipolar recoding is obtained by introducing the exploring electrode (anode) of the recording amplifier inside the heart in contact with the myocardium. Fig. 2.5 (A) shows that as the depolarization wavefront propagates towards the electrode, a positive deflection (R wave) is generated and when the wavefront travels away, the deflection sweeps steeply negative (S wave). The second electrode (cathode) of the amplifier is kept far away from any intracardiac activity. There are two ways to position cathodal input: 1) taking advantage of surface ECG electrodes that are already placed on the skin (right arm, left arm and left leg) by coupling the three electrodes of Wilson central terminal (Fig. 2.5 (C)); 2) placing an indifferent electrode in the body itself but remote from the heart, down the inferior vena cava. Unipolar EGMs are used for their ability to detect the early depolarization sites thus detecting early atrial and ventricular focal arrhythmias. However, this unipolar recording is not recommended for low voltage pathological tissues because it still contains significant far-field signals.

- **Bipolar**
  A Bipolar signal is generated by connecting two electrodes placed in contact with the cardiac area of interest to the recording amplifier as shown in Fig. 2.5 (B). At each time stamp, the voltage recorded is the sum of two potentials: positive and the negative inputs. However, the cathodal input is inverted thus being subtracted from the positive one. This subtraction allows to eliminate the far-field signal because this noisy information is almost similar at each instant in time, thus conserving almost only the local signal. The fact that far-field signals are significantly reduced in this type of recordings facilitates the identification of local depolarization in abnormal myocardial zones such as scars of post-infraction. The Bipolar recording is used in the localization of focal arrhythmias by identifying the earliest activation point with respect to a stable reference. The morphology of bipolar EGMs depends on the direction of the wavefront added to the position of the recording electrodes [14]. As a result, the morphology of EGMs presents different patterns such as RSR' or QRS.

2.4 Cardiac Arrhythmia

Under normal blood pumping conditions, the valves ensure blood to flow in only one direction following a sequence of highly organized contractions of the four chambers.

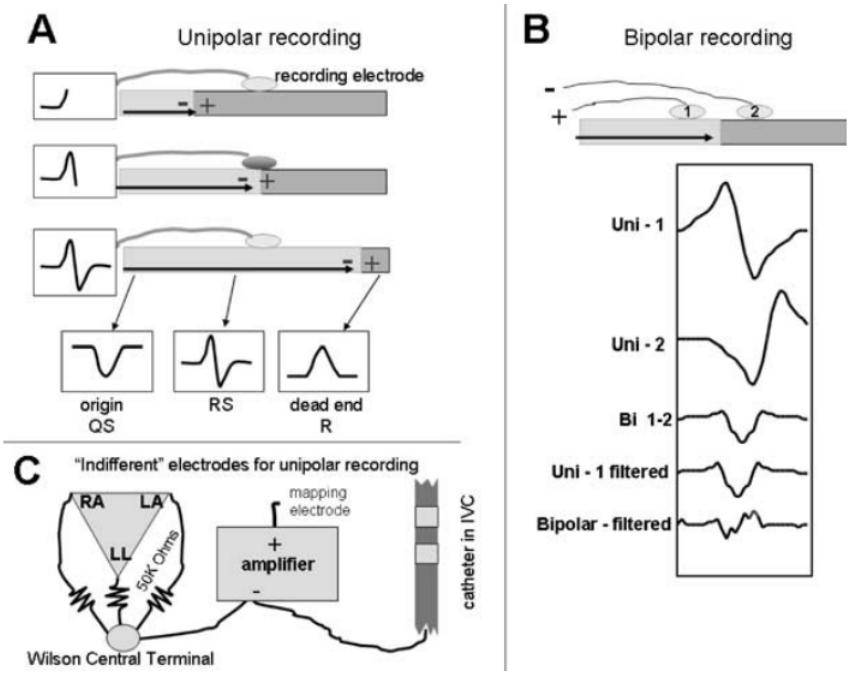

Figure 2.5: EGM recording. (**A**) Genesis of unipolar recordings. (**B**) Genesis of bipolar recordings. (**C**) Scheme of Wilson central terminal. Image taken from [13].

Therefore, any change in the sequence of electrical impulses causing a dysfunction in the blood circulation process or disturbance in the heart beating symphony reflects a heart disease, known as arrhythmia [15]. In NSR, values of heart rate vary and depend mostly on age and whether the person practices physical exercise frequently or has a sedentary life style. But a typical normal heart rate varies between 60 to 100 beats per minutes (bpm) according the American Heart Association[1]. Arrhythmias can be classified into two categories with respect to the heart rate:

- **Bradycardia** stands for a slow heart rate of less than 60 bpm. The arrhythmia class is driven by electrical impulses that are not formed in the heart's natural pacemaker (SA node) or that are not sent to the ventricles through the proper channels.
- **Tachycardia** is catheterized by fast heart rate of more than 100 bpm. Depending on the starting position of the rapid heart rate, this cardiac condition is said to be ventricular (if it begins in the ventricles) and supraventricular tachycardia, also known as atrial arrhythmia (if it begins above the ventricles).

Atrial flutter, junctional rhythm, atrial fibrillation and ventricular tachycardia are common kinds of arrhythmia, mainly among the elderly. Common factors for devel-

[1] https://www.heart.org/en/health-topics/high-blood-pressure/the-facts-about-high-blood-pressure/all-about-heart-rate-pulse

oping arrhythmias are sedentary life style, obesity, alcoholism, tobacco, caffeine excess, some over-the-counter and prescribed medicines, illegal recreational drugs and damaged heart tissue because of an illness such as stroke or heart attack. Also a wide range of studies prove the link between arrhythmias and heredity and genetics [16].

2.5 Atrial Fibrillation

2.5.1 Presentation and Statistics

Atrial fibrillation, AF, is considered as the most frequent sustained arrhythmia experienced in clinical practice and it rises in prevalence with advancing age. Assimilated to a new epidemic, AF is estimated to affect 6-12 million people in the United states (US) by 2050 and 17.9 million people in Europe by 2060 [2]. This complex tachycardia is associated with a five-fold stroke risk increase that is likely to be more severe than in non AF-related stroke [3] [5]. Added to that, AF is related to a two-fold increased risk of both mortality and dementia. Moreover, hospitalizations with AF as the primary diagnosis represent nearly half million annually and this disease is estimated to contribute to almost one-hundred thousand deaths per year in the US alone. As a result, AF provokes a significant morbidity and mortality risks accompanied with an important economic burden. Hence, studying this complex arrhythmia and developing decision-aid solutions for faster and safer diagnosis of AF, with the use of signal processing and ML tools, presents a clear socio-economic interest. Some major causes of AF are due to rheumatic disease, atherosclerotic disease, hyperthyroidism or rarely to strong sympathetic activation in healthy subjects.

2.5.2 Mechanisms

The genesis of a normal heart rhythm is described earlier in this chapter. It is characterized by a synchronous activation between the atria and the ventricles. NSR happens when the SA node initiates an electrical impulse that travels towards the left atria (LA) and the ventricles. However, AF occurs when there exist a fully irregular and chaotic activation in the atria, typically around the pulmonary veins. Then instead of beating effectively to eject blood into the ventricles, the heart starts quivering or fibrillating, thus causing non-synchronous fluctuations in the baseline. AF is a tachycardia because the ventricular rate becomes more rapid and irregular. The disorganized electrical impulses lead to chaotic atrial contraction, causing the atria to fibrillate, thus yielding an incomplete blood ejection towards the ventricles and blood stagnation in the atria. This may lead to strokes. Figure 2.6 shows that if a clot is formed due to blood stagnation, it may

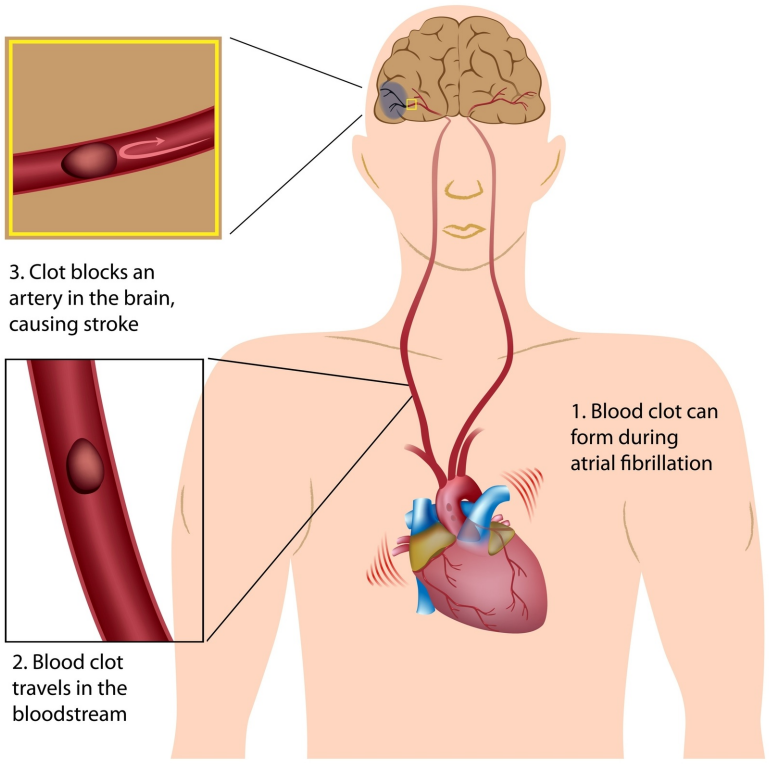

Figure 2.6: Blood clot formation in AF[2].

cause a stroke.

The atrial EGM of subjects suffering from AF is different from the NSR. It is characterized by the a signal fractionation [7] added to a significant delay in intracardiac activations [9]. The AF cycle (AFCL) length is the time between two successive atrial depolarizations.

The specific electrophysiological mechanisms leading to AF are still partially undiscovered regarding the complexity of this arrhythmia [17]. The three prevailing mechanisms include electrical, structural and automatic remodeling.

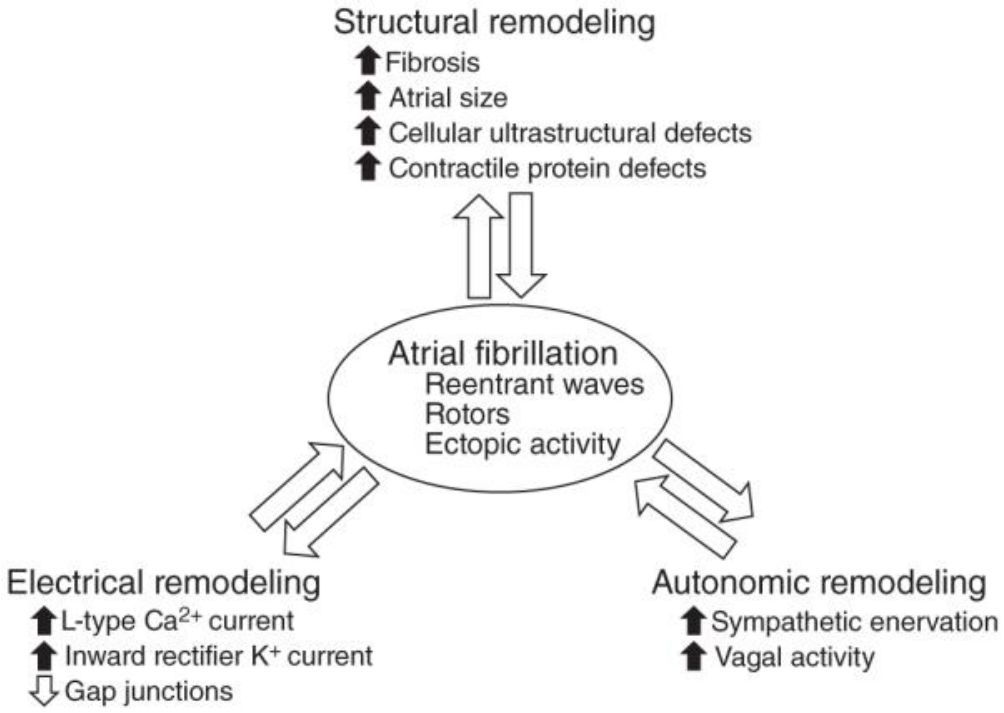

Figure 2.7: Diagram representing the major types of remodeling that lead to AF. Image taken from [17].

- **Electrical Remodeling**: This remodeling occurs in atrial cardiomyocytes (cells of cardiac muscle) [18]. A key advance in understanding the mechanisms of AF consists in identifying the molecular determinants of abnormal atrial electrical activity like ion channels, gap junctions and regulators of intracellular Ca^{2+} homeostasis. Studies on animal models and clinical research helped identify several categories of ionic currents that change in AF and contribute to its development such as L-type Ca^{2+} and inward rectifier K^{+} currents. The change in the gap junction function is correlated with conduction velocity, which is a well-known determi-

[2] https://www.bloodclot.org/atrial_fibrillation

nant of the arrhythmia. Indeed, the conduction velocity gets slower and favors reentry, allowing for initiation and maintenance of AF.

- **Structural Remodeling**: Structural remodeling is related to changes in the atria, characterized by transformations in atrial tissue properties into fibrosis, changes in atrial size and cellular ultrastructure. As a result, the atria defects in conduction thus contributing to reentry and rotor formation.

- **Autonomic Remodeling**: The cardiac electrophysiology is significantly controlled by the autonomic nervous system. Indeed, the heart is a massively innervated organ with both extrinsic (ganglia outside the heart) and intrinsic (ganglia inside the heart) nervous tissues. The vagal nerve and nerves arising from the paravertebral ganglion, forming the extrinsic nerves, are affected. In this context, AF has been associated with defects in the cardiac autonomic function such as increased sympathetic activity.

2.5.3 Maintenance Factors

Even though the major factors underlying the maintenance of AF are still unclear, the literature has highlighted the following theories:

- **Reentry circuits** also called an accessory pathways or multiwavelet circuits happens when two different APs, the first with fast conduction and long refractory period and the second with slow conduction and short refractory period, travel a common conduction tissue to form a circuit.

- **Triggers** are typically located around the pulmonary veins (PV) and are related to the initiation of AF. The stepwise ablation targets the triggers by isolating the PVs [8].

- **Substrate** or gradient refers to atrial sites with damaged conduction and non-synchronous electrical activations that are held to be responsible for maintaining and self-perpetuating AF [19] [9].

- **Autonomic sources** represent sites with a great number of autonomic nerves, present throughout the atria, that may play the role of AF sources, as a consequence of autonomic remodeling explained in Sec. 2.5.2

- **Rotors** are basically defined as spiral waves at high speed into surrounding tissues. Clinically, a rotor is characterized by repetitive and cyclic activation around a core.

2.5.4 Stages

The malignancy degrees of arrhythmias differs according to the appearance rate of abnormality episodes and its duration. AF evolves gradually from short to longer episodes and may degenerate into sustained forms that are characterized by frequent attacks. However, regression of AF from persistent to paroxysmal is also possible. Depending on episode duration and response to treatment, AF can be classified into five stages based on the clinical guidelines of the European Society of Cardiology [4]:

- **First diagnosed AF**, if first diagnosis of an AF episode without taking into account the duration nor the gravity of AF symptoms.
- **Paroxysmal AF** represents intermittent AF episodes as it commonly self-terminates in 48 hours and does not exceed 7 days.
- **Persistent AF** this form of AF is complex as it lasts more than 7 days and is generally terminated with cardioversion. The drivers perpetuating persistent AF are particularly complex.
- **Long-standing persistent AF** where persistent AF presents longer and more continuous episodes that may last for more than one year. Therapies to restore the normal heart rhythm become necessary.
- **Permanent AF** when protocols to restore the normal heart rhythm fail and when AF is present in a permanent manner. AF becomes accepted by both patient and electrophysiologist. Rate control therapies are put in place.

2.5.5 Therapies

Given the complexity of understanding the exact mechanisms deriving AF at each stage, treatment methods are still less than optimal and represent a rich subject of ongoing research and clinical experimentation. Treatments of AF differ according to several factors, including AF episode duration and gravity, age, patient's overall health, AF-related symptoms, whether the subject has an underlying cause that needs to be treated and mainly the type of AF. Among the exiting therapies of AF, pharmacological treatment, external electrical cardioversion (EEC), catheter ablation (CA) and surgery represent the four pillars of curing or alleviating FA, as summarized next:

- **Pharmacotherapy**:
  Drug therapy is adapted to early stages of the arrhythmia like the paroxysmal type where AF episodes last less than 48 hours. If AF symptoms persist longer, pharmacology becomes less effective [20]. In order to reduce the risk of developing blood

clots that may potentially cause stokes, the patient may be prescribed anticoagulants like dabigatran or rivaroxaban. There are also medicines to control both heart rate and heart rhythm. Drugs like atenolol, verapamil and propranolol tend to slow the fast heart rate in this tachycardia. But they do not correct the irregular rhythm or pattern of AF. In order to restore the NSR and maintain a regular heart rhythm, damaged by the non synchronization between the atria and ventricles in AF, it is recommended to use beta blockers and flecainide. Similarly to any medicine, the previous anti-arrhythmic drugs can cause side effects. The most common undesired symptoms include nausea, tiredness and low blood pressure.

- **EEC**:
  This procedure consists in stimulating the heart by sending controlled electric shocks with the use of electrodes attached to the chest of the subject. The goal is to restore NSR when medication is not efficient [21]. The electric current is low-voltage and acts on the totality of the heart cells to make it perform contraction in a synchronous way. The electrical stimulation can be repeated in case the first shocks are not enough the restore the abnormal heart rhythm. The history and frequency of AF episodes should be monitored before applying the cardioversion, otherwise it could increase the risk of forming a clot.

- **CA**:
  During this invasive procedure, the cardiologists threads long and thin tubes called catheters into the heart, then applies heat (RF ablation) or extreme cold (cryoablation). This causes tiny scars in certain parts of the heart muscle, which disrupt or eliminate the erratic electrical signals in the heart. Ablation operates on heart tissues to restore NSR by destroying directly the abnormality sources (triggers or drivers). The choice of the ablation therapy depends on the cause driving the irregular heart rhythm and whether the patient presents other heart conditions. The main types of CA are the following [6]:

  - **RF Energy**: An increasingly used therapy consists in burning the cardiac myocytes (cells) displaying irregularities with RF energy delivered through a catheter. This procedure includes classical methods like the stepwise or more recent methods based of STD pattern for instance. We would like to highlight that STD-based ablation will be the focus of the present doctoral thesis.
  - **Cryoablation**: This process uses extreme cold to destroy diseased heart tissues. Cryoablation uses hollow needles, called cryoprobes, through which circulate cooled fluids. Similarly to catheters with RF energy, cryoprobes are placed adjacent to the target sites in the heart to destroy the diseased tissues.

- **Open Heart Maze Procedure**:
  Maze is an invasive procedure performed during open-heart surgery [22]. The cardiologists create a pattern of scar tissue (maze) in the atria using a scalpel to block the abnormal impulses. Electrical barriers are then created within the atria to prevent macroentrant rhythms. The scarring disrupts the stray electrical signals responsible for maintaining the arrhythmia. This surgical treatment is only used when catheter ablation is not effective or in case another heart surgery is needed such as valve repair.

2.6 Summary

The heart is a muscular pump responsible for blood circulation in the body. It irrigates the different organs with oxygenated blood and vehicles the used blood to pulmonary system in order to pure, oxygenate and feed the body with nutriments. This pumping function translates into electrical activity that can be measured through ECG and EGM recordings. The EGM characteristics of normal and abnormal heart have been presented in this chapter. We explained the mechanisms underlying AF and the main therapies. It was also highlighted that the exact factors of this complex arrhythmia are still unclear which explains the uncertain success rate of existing therapies. The next chapter will provide further details about CA of AF with RF energy and will introduce STD-guided ablation therapy.

Chapter 3

AF Ablation Guided by Spatiotemporal Dispersion

Contents

3.1 Introduction

The preceding chapter presented AF and introduced its main therapies. Because of its longterm effectiveness, CA is becoming increasingly used to treat persistent forms of AF and represents the focus of the thesis. Developing ML tools for guiding STD-based ablation is the main objective of this thesis. The present chapter provides further details about CA, starting by introducing cardiac mapping systems and technologies. High-resolution multieletrode catheters, mainly the PentaRay, are introduced and the state-of-the-art about CA with RF energy protocols of AF is drawn. The definition of

STD pattern is provided after that STD-based ablation therapy is explained. Finally, we define the endpoints of AF ablation.

3.2 Cardiac Mapping

Before ablation, an electrophysiology (EP) study is systematically performed. It involves introducing diagnostic catheters inside the heart chambers and mapping how electrical signals propagate through the heart of subjects affected with the arrhythmia [23]. Cardiac mapping consists in reconstructing a three-dimensional (3D) geometrical representation of the heart chambers with measures of electrical intracardiac potentials. There exists two main classes of mapping systems: sequential and continuous. For each mapping type, a suited high-resolution exploratory catheter is associated, such as the PentaRay, Basket and Lasso [24] [25].

3.2.1 Sequential Mapping

Electroanatomical mapping represents the principal sequential mapping technology. A magnetic field is used to reconstruct 3D maps of the heart chamber explored during the mapping. CARTO system (Biosense Webster Inc, Irvine, CA, USA) and the Basket technology use this technology:

- **CARTO mapping**: First, the endocardial geometry is reconstructed in 3D space. Second, activation times, locations and EGM signals (unipolar and bipolar) are collected and displayed. The 3D representation has a color code to distinguish the different types of electrical activations. A voltage map showing the peak-to-peak EGM amplitudes sampled at each mapped site is also provided. Therefore, the mapping process is time consuming because it is governed by the amount of mapped sites. Real-Time Position Management is another example of electroanatomical mapping that uses ultrasound technology [23].
- **Basket mapping**: The Rhythmia mapping system uses the basket catheter (Boston Scientific Inc, Marlborough, MA, USA) to provides a quick and precise 3D continuous mapping [26] [27]. It has an eight-spline basket design and 64 flat and low noise electrodes of a surface of 4 mm^2, as shown in Fig. 3.1. This high resolution catheter has bidirectional steering ability. The basket catheter is used to position atrial sites displaying rotors [28].

Figure 3.1: Basket mapping catheter (image from Boston Scientific Inc).

3.2.2 Continuous Mapping

Continuous mapping systems allow to characterize the cardiac rhythm in few heart beats (one to two) thus creating accurate 3D intracardiac maps quickly. The Abbot Topera Medical mapping and non-contact catheter mapping are two examples of continuous mapping:

- **Non-contact Technology**: Contrarily to basket technology where the electrode must be in contact with the chamber's walls, non-contact mapping is introduced transvenously inside the blood pool of the chamber to be explored. It can remotely measure the potentials generated by far-field EGMs, then an inverse problem is solved to reconstruct the EGMs [29].

3.2.3 PentaRay Catheter

The PentaRay is a high-density multielectrode (20-pole) mapping catheter (Biosense Webster Inc, Irvine, CA, USA) designed to create high definition and precise anatomical maps quickly and with a minimal number of maneuvers required to reposition the catheter. The PentaRay has a five-branch star design with 5 electrodes (2 bipoles) on each spline as shown in Fig. 3.2 [30] [31] [32]. Splines are soft and flexible. There exists four different references for this catheter with different dimensions but we are interested in the present doctoral thesis in the PentaRay used in CHU of Nice for STD-guided ablation of AF. The diameter of a branch is 3 Fr (French). Fr is a common measure unit of catheters' external diameter and represents three times the diameter in millimeters. Each electrode has a 1 mm length. The edge-to-edge inter-electrode space is 2 mm - 6 mm - 2 mm, as shown in Fig. 3.3. The PentaRay catheter associated with CARTO system allows to sample numerous points of cardiac mapping data simultaneously with reduced procedure and fluoroscopy time. This catheter maps signals precisely and validates termination of abnormal signals in AF (CFAE and STD), ventricular and atrial tachycardias (AT) [30].

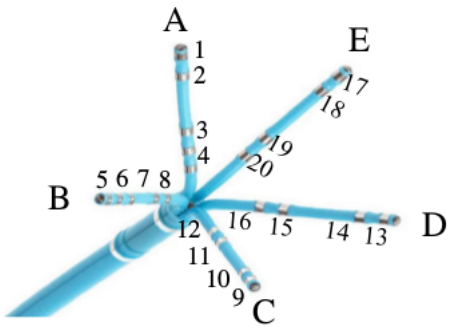

Figure 3.2: PentaRay multi-spline catheter (modified from BioSense Webster Inc).

3.3 CA of AF with RF Energy

3.3.1 Overview

Recent guidelines recommend CA for treating symptomatic AF subjects (paroxysmal and persistent) as an important rhythm-control therapy [33] [34]. Indeed, data from large studies point out that this strategy might reduce risks of mortality, heart failure and stroke.

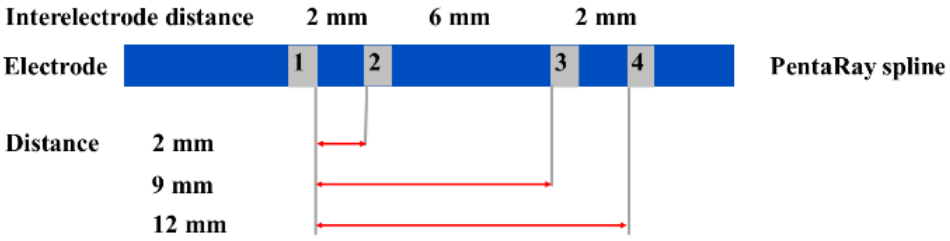

Figure 3.3: PentaRay dimensions (modified from BioSense Webster Inc).

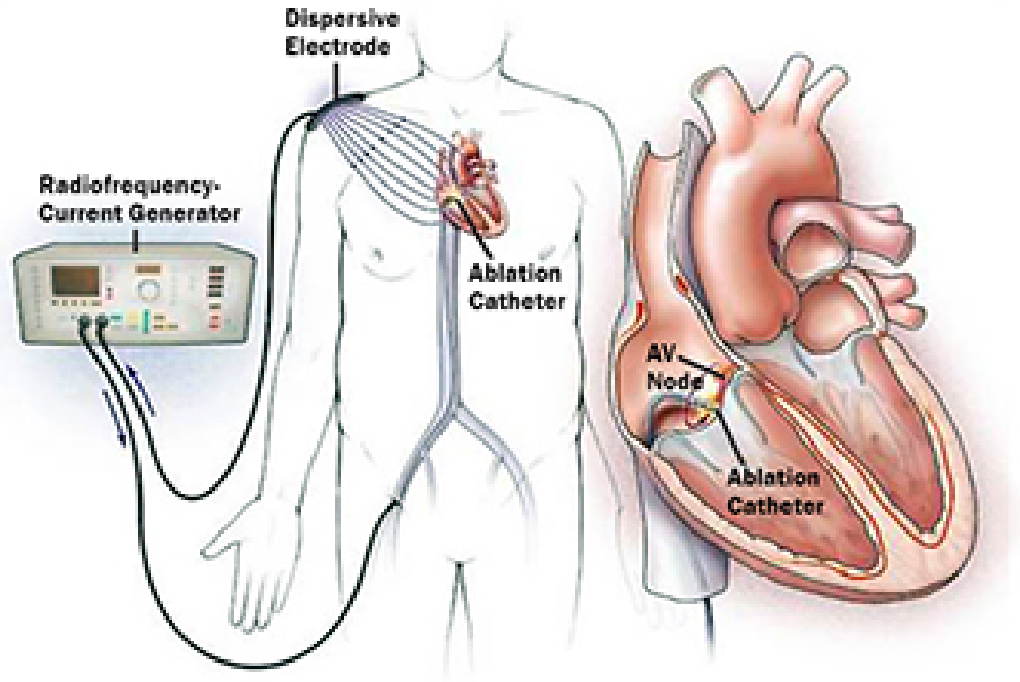

Figure 3.4: CA with RF Energy[1].

Fig. 3.4 illustrates the process of CA with RF energy. This procedure uses mapping technologies to reconstruct the EP of the heart thus tagging arrhythmogenic sites inside the chambers' walls to be ablated with the ablating catheter in the second phase of the procedure, as detailed in Sec. 2.5.5. During the last decade, several technological advances have been developed as an attempt to enhance the efficiency of AF ablation in terms of:

- Procedural approaches: disconnecting PVs, left atria (LA) linear ablation, fibrosis and dispersion based ablation.
- Tools: contact force catheters.
- Technologies: high-density, high-resolution 3D mapping, rotor mapping

The first study revealing that the PVs represent an important source of triggering foci for paroxysmal AF was performed in 1998 [35]. PV isolation (PVI) using point-to-point RF represents the cornerstone technique of CA of AF, as highlighted in [34]. Ablating atrial tissues displaying complex fractionated electrograms (CFAE), whose definition is provided below, represents also a common practice [7].

[1] https://www.webmd.com/heart-disease/atrial-fibrillationhttps://www.webmd.com/heart-disease/atrial-fibrillation

CFAE

> **Complex Fractionated Electrogram**
>
> *"CFAEs are defined as low voltage atrial electrograms (ranging from 0.04 to 0.25 mV) that have fractionated electrograms composed of two deflections or more, and/or have a perturbation of the baseline with continuous deflection of a prolonged activation complex. CFAE has a very short cycle length (≤120 ms) with or without multiple potentials; however, when compared to the rest of the atria, this site has the shortest cycle length." [33]*

CFAE signals are defined as atrial electrograms that have fractionated morphologies composed of several deflections and are characterized by very short cycle lengths. In [7], the authors provide evidence that CFAE areas are critical sites for AF perpetuation and may serve as target sites for AF ablation. However, several studies show the limitations of ablating CFAE sites alone added to the non-reproducibly of successful results in CFAE-based ablation. This approach uses a single electrode that is not enough for rotor visualization. Fig. 3.5 shows different examples of EGM signals. Grade 1 illustrates an EGM with CFAE pattern and grade 5 represents a normal one, as reported in [37]:

- **Grade 1**: Uninterrupted fractionated activity characterized by fractionated activity spreading over $\geq$ 70% of the sample. It contains at least one uninterrupted episode of fractionation of more than one second.
- **Grade 2**: Interrupted fractionated activity occupies $\geq$ 70% of the sample.
- **Grade 3**: Intermittent fractionated activity where the fractionation pattern occupies 30% to 70% of the sample.
- **Grade 4**: Complex EGM contain $\geq$ 5 direction changes and represents a discrete EGM ($<$ 70 ms). It also contains fractionated activity occupying $<$ 30% of the sample.
- **Grade 5**: Normal EGM is both discrete and simple ($\geq$ 4 direction changes).
- **Grade 6**: Scar or absence of electrical activity.

In order to optimize the outcome of CA, combinations of different ablation strategies are incorporated in order to handle the heterogeneous mechanisms underlying AF (triggers, rotors, substrate, reentry). A classical sequential ablation approach is called stepwise, it targets CFAE after having eliminated the triggers around the PVs [36]. Recent ablation strategies have also emerged such as targeting STD EGMs as potential AF drivers.

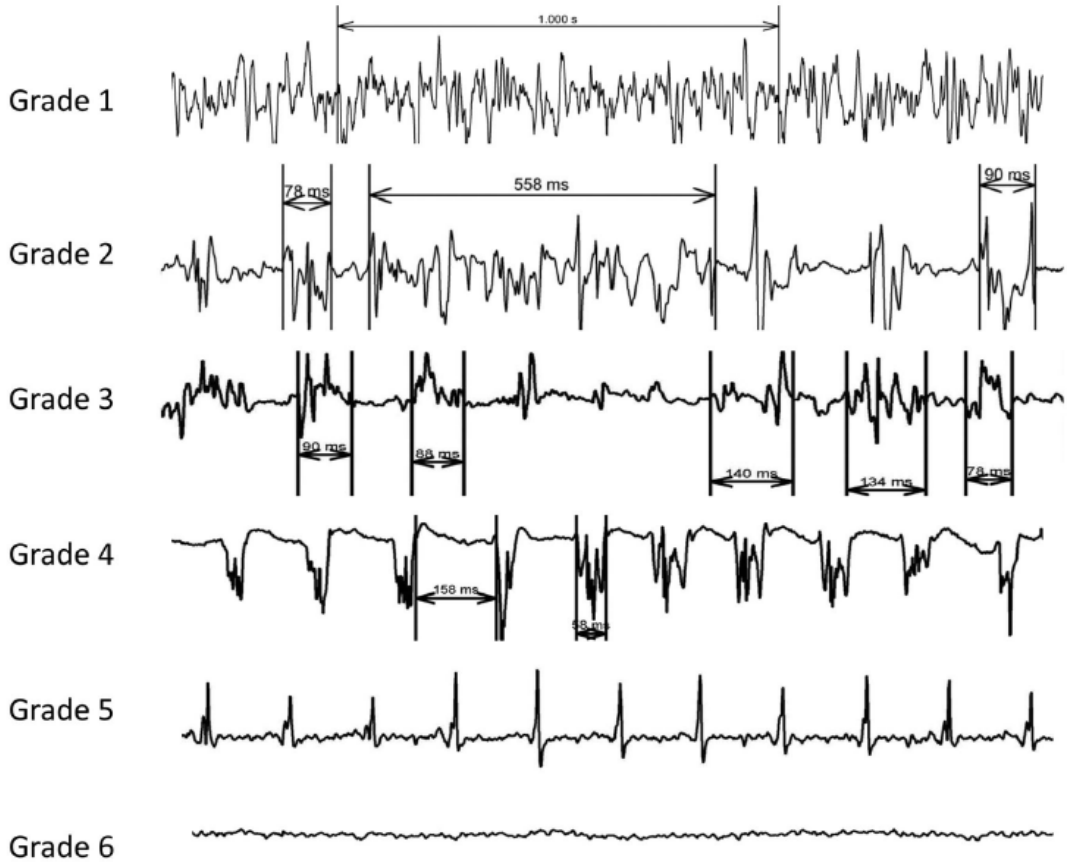

Figure 3.5: Classification of EGMs and illustartion of CFAE. Image taken from [37].

Endpoints of AF ablation

The efficacy of CA by RF energy is evaluated on both short and long terms [9]. The short-term procedural endpoints are categorized into:

1. Primary: AF termination, the AF may turn ideally into NSR or other stable AT that are less dangerous and more controllable like Atrial Flutter.

2. Secondary: conversion to NSR without external EEC and post-ablation AF non-inducibility.

The long-term endpoints concern the non reproducibility of the arrhythmia but need a representative follow-up period:

1. Primary long-term: AF recurrence rate after a single procedure, with or without anti-arrhythmic drug therapy.

2. Secondary: Related to AF recurrence rates with or without therapy.

3.3.2 Stepwise Ablation

In the classical sequential stepwise approach [38], interventional cardiologists perform three steps. First, isolate electrically the PVs (PVI) thus ablating the triggers around the PVs responsible for initiating AF. Second, the left atrial roof line is ablated. Third, atrial areas displaying CFAE are ablated using a bipolar mapping catheter. This ablation strategy allows to terminate long-lasting persistent AF in 87% of patients and achieves freedom from in 95% of patients as reported in [39]. However, a growing number of reports find little advantage in targeting CFAEs after PVI for persistent AF [9] [8]. Other studies still confirm that ablating *good* CFAE locations is effective and allows to terminate AF without having to perform routinely PVI [40]. It is suggested that the unsuccessful randomized studies claiming that ablating CFAE does not yield substrate termination, such as [8], might not be selecting the right CFAE sites and recognizing that all CFAE patterns are not the same. Instead, recent ablation strategies suggest to ablate potential AF *drivers* as they are responsible for maintaining and self-perpetuating AF [9]. The drivers are reported as atrial sites displaying STD intracardiac activations.

3.3.3 STD-Based Ablation

A recent discovery in AF ablation therapy suggests that targeting only cardiac areas displaying STD EGMs can terminate AF in 95% of cases (procedural success rate) in a cohort of 105 patients [9]. The resulting recurrence rate within 18 months of follow-up is equal to 15%. STD is a visually discernible AF pattern that guides interventional cardiologists in ablating persistent AF drivers. STD-based ablation uses PentaRay high resolution mapping catheter. Before ablation with RF energy, interventional cardiologists first position sequentially the PentaRay in different sites of the right (RA) and left (LA) atria. Ten bipoles are then simultaneously recorded per location by maintaining the catheter stable for at least 2.5 s, more than ten times a typical AFCL. The mapped locations (points) exhibiting a non-synchronous cardiac activity observable on a minimum of three contiguous bipoles are said to be "STD" points and are tagged for ablation. Hence, this ablation protocol is a wholly patient-tailored procedure where the STD pattern is considered as a visually recognizable electric footprint of AF.

In the framework of the present doctoral thesis, we study signals acquired from ablation interventions applying PVI followed by the ablation of atrial sites displaying dispersion EGMs. A recent similar study demonstrates that STD-guided ablation combined with PVI yields high rates of AF termination [41]. The study is conducted on a population of 44 patients aged 68 years in average and presenting several comorbidity factors (hypertension 89%, diabetes 21%), obstructive sleep apnea 37% and coronary artery disease 26%). The short term endpoints are satisfying. Procedural termination of AF is

observed for 73% of the subjects and conversion to NSR occurs in 38% of the study population. After a 6-months follow-up, the recurrence rates of AF (14%) and of AT (20%) are low. The notion of spatiotemporal pattern appears in [42] as an assessment criteria of the atrial activity organization during the ablation of persistent AF.

STD EGMs

The definition of STD EGMs is the following:

> **Spatiotemporal Dispersion**
>
> "*Dispersion areas were defined as clusters of electrograms, either fractionated or nonfractionated, that displayed interelectrode time and space dispersion at a minimum of 3 adjacent bipoles such that activation spread over all the AFCL.*" [9]

According to guidelines in [9], dispersion areas refer to clusters of electrograms, either fractionated (CFAE) or not, displaying interelectrode time and space dispersion at a minimum of three contiguous leads, as shown in Fig. 3.6 (A). The authors of that study claim that the presence of an underlying source of AF is recognizable through the fractionation occurring in a non simultaneous fashion at neighboring electrode locations (time dispersion) and organized in well-defined clusters (spatial dispersion) as illustrated in Fig. 3.6 (A). The dispersion is visually recognizable and can be quantified by a 70% delay relative to AF in the cardiac activation of a location compared to its neighborhood. For an AFCL of 200 ms, if the cycle is gradually delayed through three contiguous bipoles by 140 ms then the delay is $140/200 = 70\%$. Fig. 3.8 illustrates dispersion. At each bipole in a dispersion area, electrograms may either display or not signal fractionation. Fractionation stands for a continuous and slow activations with small amplitude. Several EGM fractionation morphologies are presented in Fig. 3.6 (B), namely: 1) continuous, low voltage fractionated EGMs; 2) bursts of fractionated EGMs; 3) fast non-fractionated EGMs; 4) slow non-fractionated EGMs. Representative examples of multipolar EGMs (Fig. 3.7) illustrate that fractionated signals are found in both dispersion and non-dispersion regions.

Limitations of Visual Inspection

It is worth highlighting that the visual inspection of STD is difficult because the delay in cardiac activation must be quantified visually by the eye of the interventional cardiologist (70% of the AFCL). We recall that thousands of locations are mapped with the PentaRay catheter in a typical ablation intervention and that all multichannel EGM recordings acquired at these locations are evaluated visually by cardiologists in real time. Hence, the visual identification of STD areas is prone to errors and lack of reproducibil-

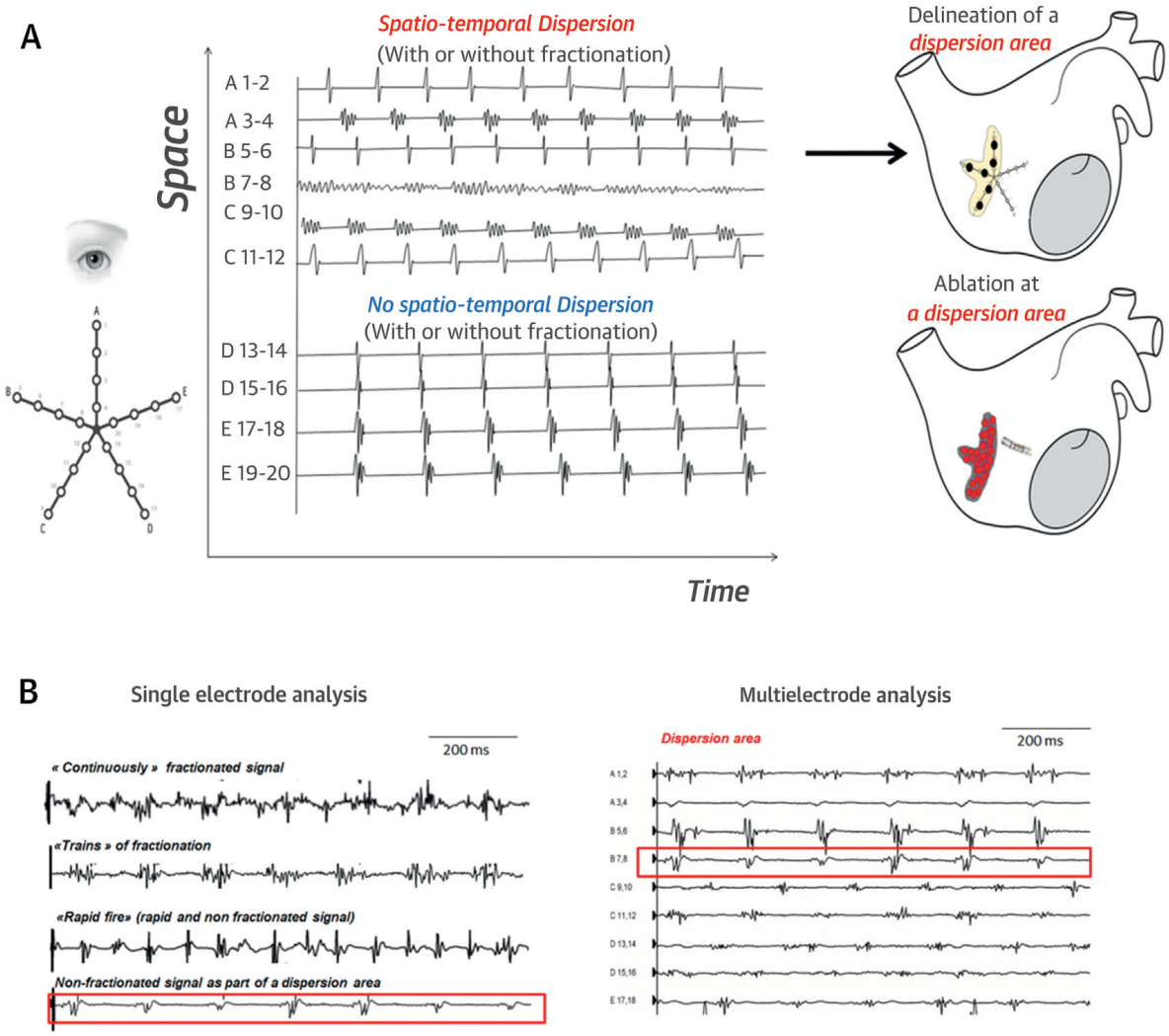

Figure 3.6: (A) Dispersion areas are defined and delineated via a mapping approach. Channels A 1-2, A 3-4, B 5-6, B 7-8,C 9-10 and C 11-12 display STD contrarily D13-14, D15-16, E-17-18, E19-20.(**B**) Single-bipole signals from dispersion regions were differentiated by fractionated or non-fractionated EGM morphologies; non-fractionated (>120 ms) EGMs may be 1 or several of the EGMs within a dispersion region. Collectively, the bipolar EGMs span most of the AFCL recorded in the region. Image taken from [9].

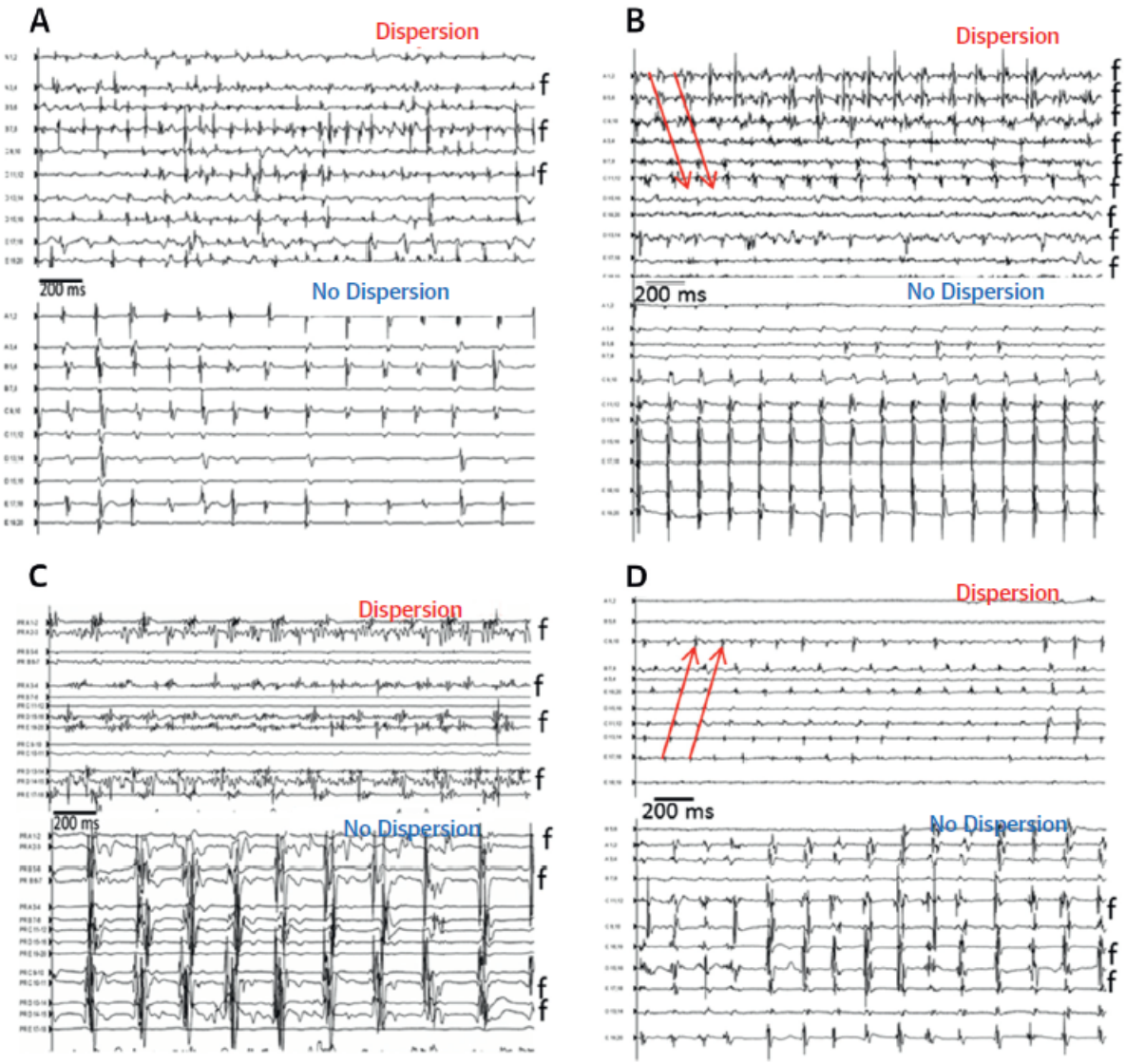

Figure 3.7: (A to D) Representative examples of multipolar EGMs recorded in dispersion and non-dispersion regions in 4 patients. In the dispersion regions, bipoles exhibiting continuously fractionated signal are common (**B and C**). **f** indicates fractionated EGMs; **red arrows** indicate sequential activation of consecutive bipoles of multielectrode catheter spanning 100% of the AFCL. Image taken from [9].

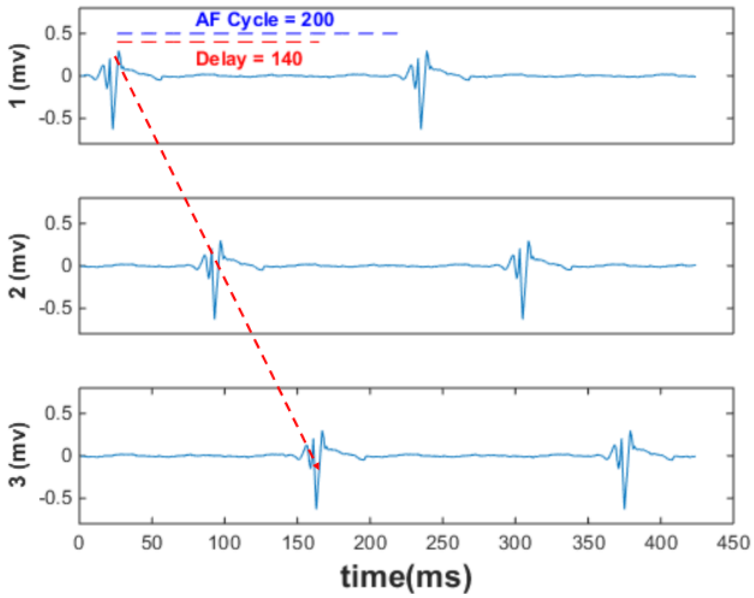

Figure 3.8: Schematic representation of the delay in cardiac activations through neighboring leads in STD.

ity and represents a time consuming task . Instead, the present work aims to design a decision-aid solution that helps interventional cardiologists detect STD pattern automatically thanks to ML tools.

3.4 Summary

The basics and recent advances in CA of AF with RF energy are detailed in this chapter. PVI is presented as a cornerstone CA strategy and several examples of CFAE electrograms are given. The stepwise ablation is introduced as a classical approach that combines PVI and the ablation of atrial sites displaying CFAE. Then STD-guided ablation is reported as a fully patient-tailored therapy that seems to outperform the classical stepwise protocol according to recently reported results [9]. STD pattern is defined as a spatiotemporal delay of cardiac activations through neighboring leads that may contain or not fractionated signals. Statistics about the success rate of STD-based ablation are also provided. Finally, we highlight the limits of the visual inspection of STD locations, motivating the need for ML-based solutions that are developed in this thesis. The next chapter will provide the mathematical background and statistical tools, necessary to the understanding of experiments conducted in the present doctoral thesis.

Chapter 4

Mathematical Background

Contents

4.1 Introduction

In the previous chapter, we presented the main AF ablation therapies and focused on the STD-guided ablation. Currently, the identification of STD EGMs is performed visually by interventional cardiologists. Instead, the present doctoral thesis proposes decision-aid solutions that automatize the identification of potential ablation sites in STD-based ablation. This chapter presents the mathematical background to understand the algorithms and statistical tools studied for designing the solutions proposed in this PhD. We start by defining conventional ML tools, then we introduce the basics of CNNs including examples of some state-of-the-art architectures used in our experiments. In addition to classification algorithms, clustering is also defined. Further statistical tools used to fine-tune and validate the classification performance are presented. Finally, we introduce the basics of linear algebra and tensor decomposition.

4.2 Machine Learning

Artificial intelligence (AI) was first introduced in the 1950s and can be defined as *the effort to automate intellectual tasks normally performed by humans* [43]. AI brings answers to questions related to whether computers could be made to conceive rational decisions. AI includes ML and deep learning (DL) [10] as shown in Fig. 4.1. Traditional program-

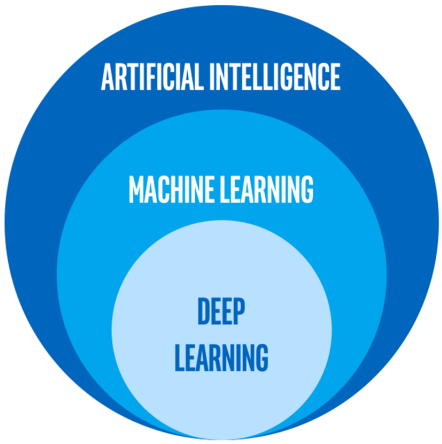

Figure 4.1: Difference between AI, ML and DL (Image from Intel Inc, CA, US)[1].

ming and classical statistical analysis require a prior knowledge of the rule that will be fed to the computer (program) along with the input data in order to deduce the answer. However, a ML program takes as input data as well as their corresponding expected answers then deduces the decision rule that will be applied to new data for which the

[1]https://www.intel.com/content/www/us/en/artificial-intelligence/posts/difference-between-ai-machine-learning-deep-learning.html

answers are unknown. The decision rule is learned from the labeled dataset and allows to produce the unknown answers. This approach is called data-driven. The difference

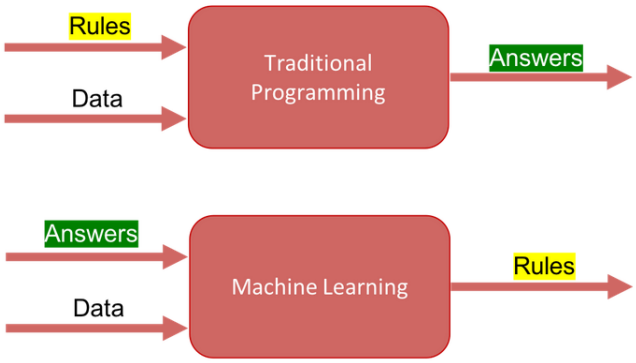

Figure 4.2: ML vs. traditional learning paradigms[2].

between traditional programming and ML is illustrated in Fig. 4.2. A ML system needs an important amount of training samples from which it will investigate the statistical structure of the input data, thus establishing the decision rule for automating the task. A ML algorithm searches automatically for a transformation that turns input data into a more meaningful and useful representation in order to make the estimated answer as close as possible to the true one, also called ground truth.

Deep learning (DL) is a subfield of ML that is related to the depth of neural networks (NN) [44]. The depth of a deep NN (DNN) represents the number of layers that contribute to the model. DL allows to learn hierarchical representations from input data to perform a task. For instance, Fig. 4.3 shows the evolution of the hierarchical data

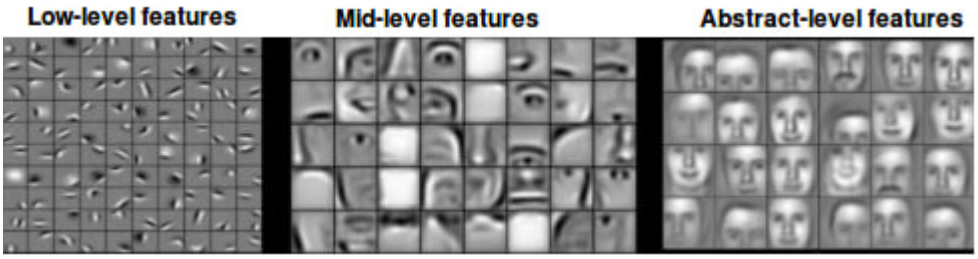

Figure 4.3: Deep representation learned by a face-recognition model[3].

representations learned by the successive layers of a deep NN for the task of recognizing human faces. As layers become deeper, the learned features become more meaningful. The ML tasks include classification, prediction and clustering.

[2] https://datalya.com/blog/machine-learning/machine-learning-vs-traditional-programming-paradigm
[3] https://www.datacamp.com/community/tutorials/machine-deep-learning

4.3 Clustering

In ML, clustering or cluster analysis is an unsupervised approach that groups a set of data points into separated and natural categories, called clusters, based on similarity criteria [45] [46]. Samples belonging to the same cluster should present a low variance and high compactness as opposite to samples from different clusters that should be as different and dissimilar as possible. An example of a clustering task consists in grouping customers according to their consumption behavior in order to propose more personalized offers. Clustering can be achieved by a variety of algorithms that differ in their definition of the similarity criteria and the distance metrics used to cluster the input objects. The most commonly used clustering algorithms are k-means [47] and hierarchical clustering analysis (HCA) [48], which are reviewed in the following section:

4.3.1 k-means

The k-means algorithm is an efficient tool to cluster data into a set of k homogeneous groups, concentrated and well separated from each other. It aims at minimizing the intra-class variances. Parting from a number of classes fixed by the user, the k-means algorithm performs as follows:

1. Assign the cluster centroids to random points in the input set.
2. Calculate the intensity-based distance from each point to the centers and assign each point to the cluster with the nearest center.
3. Recalculate cluster centers as means of each cluster.

The algorithm iterates until it converges (which occurs when no centroid is changed) and finds the optimal classification as illustrated by Fig. 4.4. Nevertheless, the problem with the k-means approach is that the result may depend on the randomly initialized centroids, and therefore the algorithm may not find the optimal clustering. Another important inconvenient is that the number of clusters k must be set by the user.

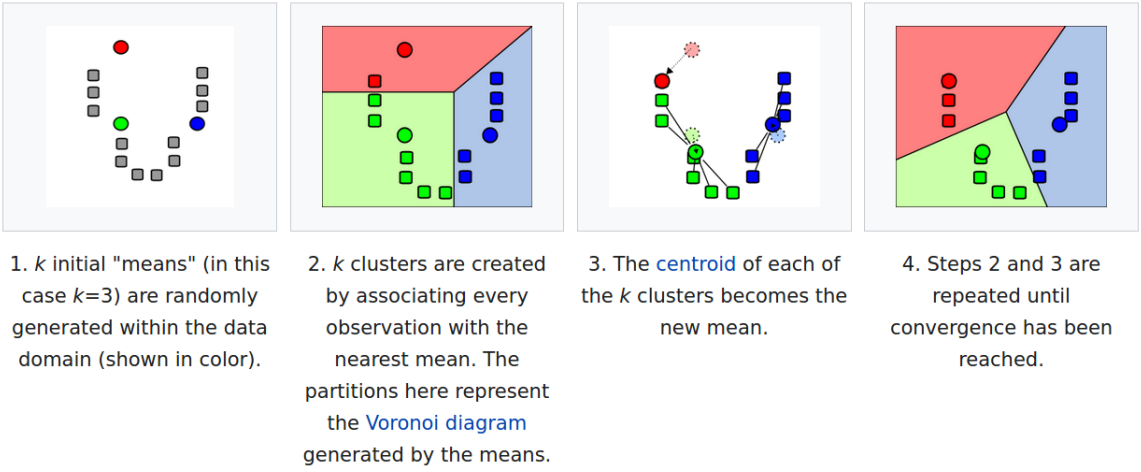

Figure 4.4: Steps of k-means standard algorithm.[4]

4.3.2 Hierarchical Clustering Analysis

Given a set of data samples, hierarchical clustering analysis [48] partitions them into agglomerations using a homogeneity criterion. This algorithm treats each observation as a singleton group at the beginning, then it merges pairs of groups until all clusters have been agglomerated into a single cluster containing all observations. The linkage can be presented by a dendrogram, as illustrated in Fig. 4.5. Each merge is represented by a horizontal line according to a linkage rule. The sim coordinate represents the similarity of two merged clusters through a similarity measure, such as neighboring or dissimilarity distances. The most important advantage of using HCA is that it does not require a previous specification of the number of clusters which is in general unknown and can drive to biasing all the clustering process if chosen subjectively, unlike the k-means algorithm [47]. In fact, after dendrogram evaluation, a cutting point, commonly called cutoff value, is specified and the clusters are formed. The choice of cutoff can be determined through numerous criteria, such as:

- Cutting the dendrogram where the gap between two successive combination similarities is the largest. Such large gaps indicate *natural* clusterings.
- Cutting at a predefined level of similarity.

4.3.3 Hellinger Distance

HCA will be used for histogram clustering in STD EGM characterization, as detailed in Chap. 7. For this reason, we opt for the distance metric of Hellinger, as it is

[4] https://brilliant.org/wiki/k-means-clustering/

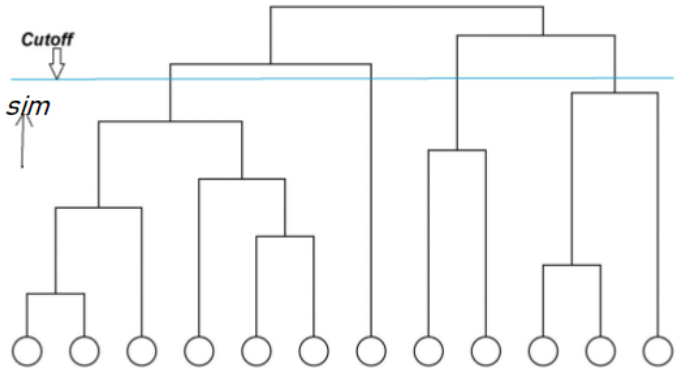

Figure 4.5: Dendrogram tree with four clusters.

adapted for sparse data such as sparse histograms. Hellinger distance [49] is a measure of similarity between two probability distributions **P** and **Q** . For discrete distributions $\mathbf{P} = (p_1, \ldots, p_k)$ and $\mathbf{Q} = (q_1, \ldots, q_k)$, their Hellinger distance equals:

$$d(\mathbf{P}, \mathbf{Q}) = \frac{1}{\sqrt{2}} \sqrt{\sum_{i=1}^{k} (\sqrt{p_i} - \sqrt{q_i})^2}.$$

4.4 Classification Algorithms

In ML, the classification task consists in predicting the class (label) of a given data point x by approximating a mapping function. The output of a classification algorithm is the predicted label ($\hat{y}$), a discrete variable. Classification is supervised as it requires a labeled dataset X to make the classification rule. When the data point x has n variables or features and can be stored in a vector ($\mathbf{x}$) of length n, $\mathbf{X}$ can be modeled as a matrix of dimensions ($m \times n$) containing the m samples $\{\mathbf{x}_i\}_{i=1}^{m}$ in its rows. This is shown in Fig. 4.6. When the data point is two-dimensional (2D image or matrix) of size ($n_1 \times n_2$), the database can be stored in a 3D array (tensor, $\mathcal{X}$) of dimensions ($n_1 \times n_2 \times m$).

In a classification task, the study dataset can be partitioned into three subsets: training, validation and test, whose definitions are the following:

- Training set contains the samples used for learning in order to fit the parameters of the classifier.
- Validation set contains the samples used to tune the parameters of a classifier, for example to choose the number of hidden units in a neural network. The validation set can be seen by the classifier during the training set contrarily to the test one.

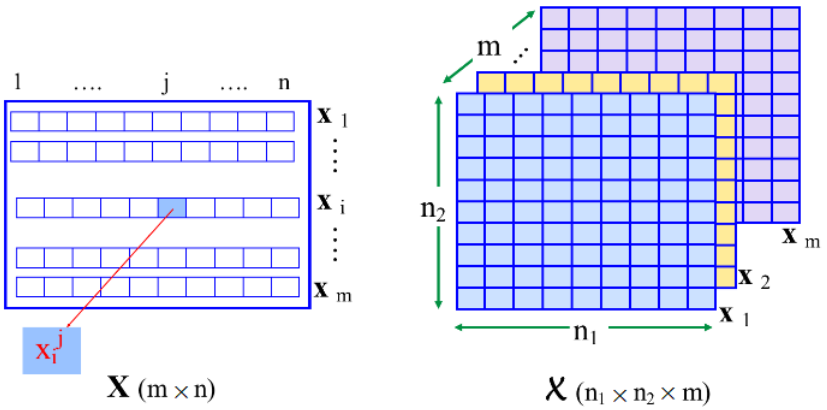

Figure 4.6: Data representation for ML tasks.

- Test set contains the samples used only to assess the performance of a fully-trained classifier.

4.4.1 Principal Component Analysis

Principal component analysis (PCA) explores linear relationships between features and defines variables significance on the basis of their variance contribution. It projects data onto a new space model formed by synthesized variables called principal components (PC). This is useful when large amounts of information may be approximated by a moderately complex model structure. Based on the estimation of the correlation structure of the variables, PCA evaluates the importance of a variable in a PC model through the size of its residual variance [50]. PCs are ordered with respect to their corresponding eigenvalues. Let $\mathbf{X}$ be the data matrix of dimensions $(m \times n)$, as explained earlier in this section. The choice to be examined is how many PC's (r) adequately account for the total variation in $\mathbf{x}$. Among the existing criteria for choosing r, we opt for the cumulative percentage of total variation, also called hard threshold (t). The optimal number of PC's is chosen as the smallest integer r for which the cumulative percentage of total variation reaches or exceeds t. Fig. 4.7 draws an example curve of cumulative variance and shows how to choose r for t =85% and t =95%, respectively.

PCA allows for variable selection and dimensionality reduction. For dimensionality reduction, PCA projects $\mathbf{X}$ onto a reduced space of dimension r. The input data is first projected $(\hat{\mathbf{X}})$ then reduced $(\hat{\mathbf{X}}_r)$:

1. Computing the eigendecomposition of the covariance matrix $\mathbf{\Sigma} = \mathbf{X}^T\mathbf{X}$. The matrix $\mathbf{W}$ is formed from the eigenvectors of Σ. Each column $\{\mathbf{w}_1 \dots \mathbf{w}_r \dots \mathbf{w}_n\}$ of $\mathbf{W}$ represents an eigenvector of $\mathbf{\Sigma}$ (new projection direction) such that $\hat{\mathbf{X}}^j = \mathbf{X}\mathbf{w}_j$, for $j = \{1, \dots, n\}$. The eigenvectors are chosen to have a unit norm and are ordered

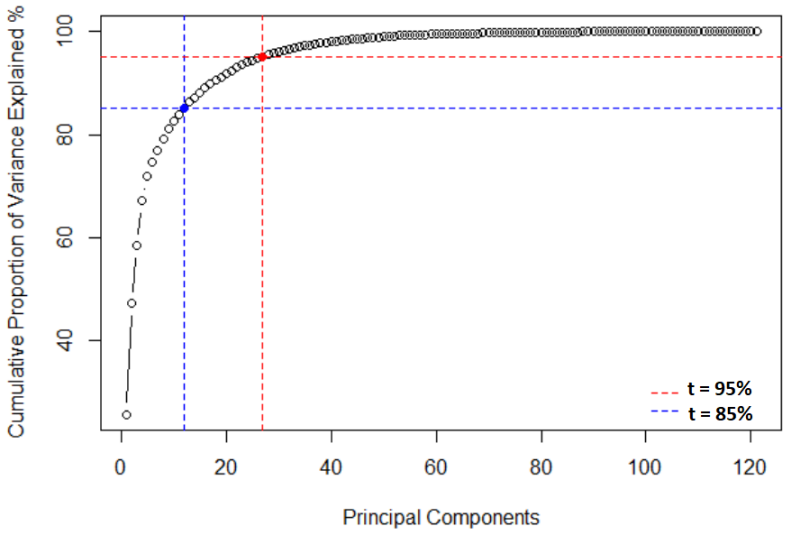

Figure 4.7: Choosing the number of PC's from the criterion of cumulative proportion of variance.

with respect to their corresponding eigenvalues, which reflects the contribution of each PC to the data variance.

2. A cutoff value t is chosen. The first r columns of $\mathbf{W}$ satisfying a cumulative variance percentage lower than or equal to t are kept to form the transformation submatrix ($\mathbf{W}_r$). The remaining eigenvectors are discarded .

3. The projected data matrix with reduced dimensionality is computed as: $\hat{\mathbf{X}}_r = \mathbf{X}\mathbf{W}_r$.

In practice PC's can directly be computed through the singular value decomposition (SVD) [51] of data matrix $\mathbf{X}$.

In the framework of our work, PCA will be used to reduce the dimensions of input data. The new data model (projected and reduced) will be fed to a classifier, as detailed in Chap. 8.

4.4.2 Linear Discriminant Analysis

The primary purpose of linear discriminant analysis (LDA) [52] [53], is to transform the input data into a new space of lower dimension which is optimal for distinguishing between groups (classes). Optimality is measured in terms of the ability of the new variables, as elements of the new LDA space, to maximize the between-class separability and minimize the within-class variability. For example, let us suppose we have a simple two variables problem in which there are two classes of objects as shown in Fig. 4.8. Each class forms a one-dimensional (1D) Gaussian distribution when projected

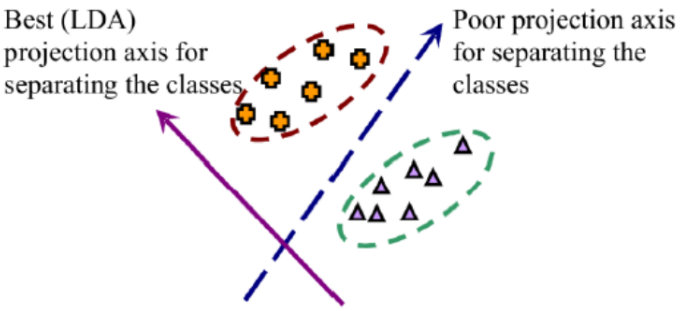

Figure 4.8: Linear Discriminant Analysis Projection[5].

on the continuous axis (purple) and can be well separated. The variance is also low on that axis, meaning that overall we can classify a new point quickly and easily. In contrast, if we were to project the two classes onto the dashed axis (blue), then the resulting distributions would overlap considerably and the classification would be very difficult. LDA allows to estimate the axes $\boldsymbol{w}_i$, $i = 0, 1$, satisfying a maximal separation between the two distributions (classes) by maximizing the between-class ($\sigma^2_{\text{between}}$) variance and minimizing the within-class variance (σ^2_{within}). Let $\boldsymbol{\mu}_0$, $\boldsymbol{\mu}_1$ and $\boldsymbol{\Sigma}_0$ and $\boldsymbol{\Sigma}_1$ denote the mean vectors and covariance matrices of the two observation classes, respectively. Then, $\boldsymbol{w} \cdot \boldsymbol{\mu}_i$ and $\boldsymbol{w}^T \boldsymbol{\Sigma}_i \boldsymbol{w}$ will correspond, respectively, to the means and variances of the linear combination of features $\boldsymbol{w} \cdot \boldsymbol{x}$. The dot product between two vectors $\mathbf{x}$ and $\mathbf{w}$ is defined as:

$$\mathbf{w} \cdot \mathbf{x} = \mathbf{w}^T \mathbf{x} = \sum_{i=1}^{n} w_i x_i.$$

As a result LDA can be formulated as the optimization of Fisher's ratio (F) [54], given by:

$$F = \frac{\sigma^2_{\text{between}}}{\sigma^2_{\text{within}}} = \frac{(\boldsymbol{w} \cdot \boldsymbol{\mu}_1 - \boldsymbol{w} \cdot \boldsymbol{\mu}_0)^2}{(\boldsymbol{w}^T \boldsymbol{\Sigma}_1 \boldsymbol{w} + \boldsymbol{w}^T \boldsymbol{\Sigma}_0 \boldsymbol{w})} = \frac{(\boldsymbol{w} \cdot (\boldsymbol{\mu_1} - \boldsymbol{\mu_0}))^2}{\boldsymbol{w}^T (\boldsymbol{\Sigma}_0 + \boldsymbol{\Sigma}_1) \boldsymbol{w}}.$$

LDA is generalizable to multiple classes and arbitrarily large numbers of variables. A recent study provides evidence that PCA can perform LDA when using ℓ_1 norm, in an unsupervised manner, thus sparing the need for labeled data [55].

4.4.3 Multivariate Logistic Regression

In statistics, logistic regression [56] estimates the probability of a given class or event by using the logistic function. Multivariate logistic regression (MLR) is identical to univariate logistic regression, but it considers more than one covariate. Let $p(\mathbf{x})$ be the model's estimation of the probability of an event that depends on n covariates or inde-

[5] https://medium.com/@srishtisawla/linear-discriminant-analysis-d38decf48105

pendent variables $\{x_1, x_2, \ldots, x_n\}$ in $\mathbf{x}$. In order to understand the mathematical formulation of logistic regression, we need to introduce the logit function. This logit corresponds to the log of the odds $(\frac{p(\mathbf{x})}{1-p(\mathbf{x})})$ and is modeled as the projection of $\mathbf{x}$ on axis $\boldsymbol{\beta}$ formed by $\{\beta_1, \beta_2, \ldots, \beta_n\}$:

$$\log(\frac{p(\mathbf{x})}{1-p(\mathbf{x})}) = \boldsymbol{\beta} \cdot \mathbf{x} + \beta_0, \tag{4.1}$$

where β_0 represents a bias. Then, the probability $p(\mathbf{x})$ is obtained by exponentiating Eq. (4.1):

$$\begin{cases} p : \mathbb{R}^n \to (0,1) \\ p(\mathbf{x}) = \frac{e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_n x_n}}{1+e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_n x_n}}. \end{cases}$$

The estimates of the unknown parameters β are derived through the maximization of a likelihood function. The sign of β_i, $i \in \{1, 2, \ldots, n\}$ indicates whether x_n will contribute to increase $(\beta > 0)$ or decrease $(\beta < 0)$ the probability.

$$\begin{cases} \beta > 0 & \to x_n = \text{risk factor in } p \\ \beta < 0 & \to x_n = \text{protection factor in } p \end{cases}$$

MLR is a baseline classifier in biomedical data analysis and electrograms classification [56]. Furthermore, MLR can be implemented as a fully connected (FC) NN.

4.4.4 Support Vector Machine

Support vector machine (SVM) is widely used in biomedical data analysis like electrocardiograms classification [57]. In the SVM model, data are divided by a hyperplane with a large gap [58] as illustrated by the example in Fig. 4.9. The algorithm aims at

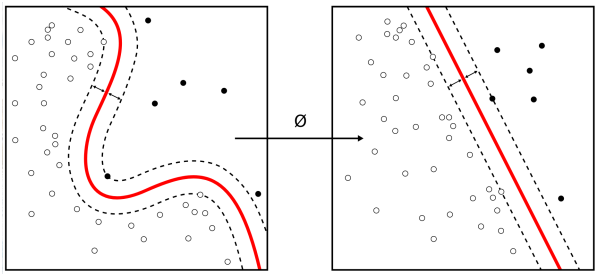

Figure 4.9: Support vector machine space. The hyperplane is plotted in red. Symbol ϕ refers to a non-linear kernel function. [6]

maximizing the distance, called margin, between the data points situated near the hyperplane, called support vectors. Maximizing the margin provides some confidence that new data points can be classified with more accuracy. Let the classification task be binary such as labels $y = -1$ and $y = 1$ correspond to samples from the negative and

positive classes, respectively. And let **w** represent a vector normal to the hyperplane, not necessarily normalized.

- In case the data are linearly separable, the width of the margin is equal to $\frac{1}{||\mathbf{w}||_2}$. So in order to maximize the margin, we need to minimize the ℓ_2 norm of **w**, denoted $||\mathbf{w}||_2$. The equation of the hyperplane is the following:

$$\mathbf{w}^T \cdot \mathbf{x}_i - b = 0,$$

  where b is a bias. Then in order to impose a large enough margin for each data point to lie on the correct side of the margin, the classification rule is computed as:

$$\begin{cases} \mathbf{w}^T \cdot \mathbf{x}_i - b \geq 1, & y_i = 1 \\ \mathbf{w}^T \cdot \mathbf{x}_i - b \leq -1, & y_i = -1 \end{cases}$$

  This constraint can be reformulated as:

$$y_i(\mathbf{w}^T \cdot \mathbf{x}_i - b) \geq 1, \text{ for all } 1 \leq i \leq m.$$

- In case data are not linearly separable, the model maps the input feature space into a higher-dimensional space so that data samples from different classes become separable. The mapping function that projects input data into this more separable space model is ϕ and it forms the kernel function $\mathcal{K}$:

$$\mathcal{K}(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i) \cdot \phi(\mathbf{x}_j) \text{ for } 1 \leq i, j \leq m,$$

  where $\mathcal{K}$ may be a linear, polynomial, Gaussian or exponential function. The Gaussian kernel, also called radial basis function is given by:

$$\mathcal{K}(\mathbf{x}_i, \mathbf{x}_j) = e^{-\frac{||\mathbf{x}_i - \mathbf{x}_j||^2}{2\sigma^2}},$$

  where σ is a hyperparameter. Looking for this suitable and separable representation of data is called the kernel trick. Then, new samples are mapped into the new space and predicted to belong to a class or another based on the side of the hyperplane on which they fall.

The optimization problem of the SVM algorithm and the different kernel choices for kernels are detailed in [59]. Contrarily to training NN that can be stuck at local minima, SVM's optimization problem is convex and yields the global maximum solution.

[6] https://www.scilab.org/tutorials/machine-learning-%E2%80%93-classification-svm

4.4.5 Neural Network

Artificial neural networks are commonly used in biomedical data analysis [56]. A NN consists of a series of connected layers. Each layer is composed of a number of artificial neurons, called nodes, that have learnable weights [60] [61].

Biological Inspiration

The animals nervous system is composed of approximately 10^{11} interconnected neurons, forming a biological neural network [62]. Each neuron consists mainly of a cell body, dendrites and an axon (Fig. 4.10). When a neuron receives signals from its neighborhood connected neurons through the dendrites, the cell body sums them. If the resulted (sum) signal is higher than a specific threshold, this electrical stimulation will be transmitted through the axon to the synapses, at chemical connections points between the axon and other neurons' dendrites. The functioning of this neural network is principally determined by its architecture and the strengths of its synapses. The biological neural network is in continuous change throughout life. As we learn new information, new parts of the neural structure can be developed by the establishment of new connections between neurons and the strength of some synapses junction can be modified. Hence, our ability to react, think and decide are stored in the neurons and are directly related to the synaptic strengths. From this description of the nervous system, an appropriate mathematical model of the neurons was built to simulate their performance. This model is called the artificial neural network, and can be trained to perform specific functions. It has been used in many fields such as automotive, banking, telecommunication and biomedicine.

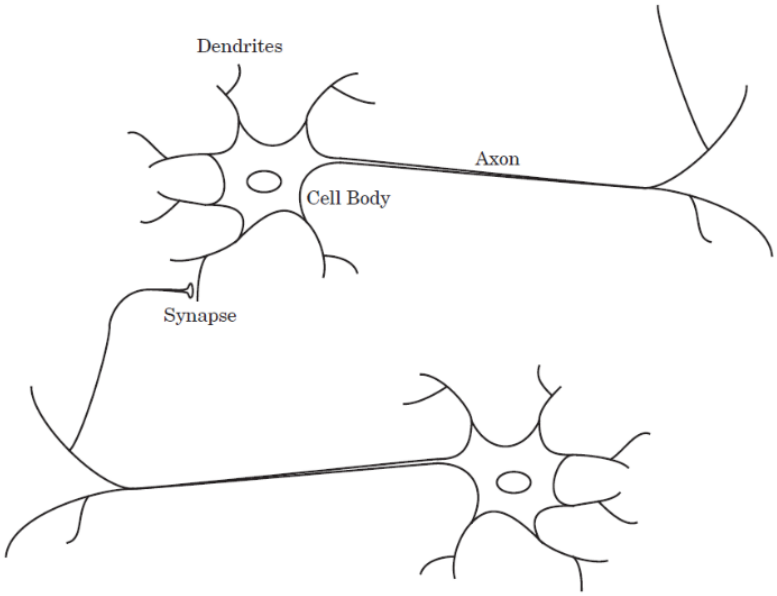

Figure 4.10: Biological neurons. Image taken from [62].

Single Input Neuron

A single input neuron is characterized by a summer Σ, a transfer/activation function f, a scalar weight w and a bias b as systematized in Fig. 4.11. The summer combined to the activation function represent the cell body in the biological neuron, while the weight corresponds to the strength of the synapse. The bias is optional, it adds an extra variable to the network which makes it more powerful. For instance, with $b \neq 0$ we can get a non null output for a null input, which is impossible without the use of b. The output of the neuron $a = f(wx+b)$ corresponds to the signal in the axon where x represents the input data.

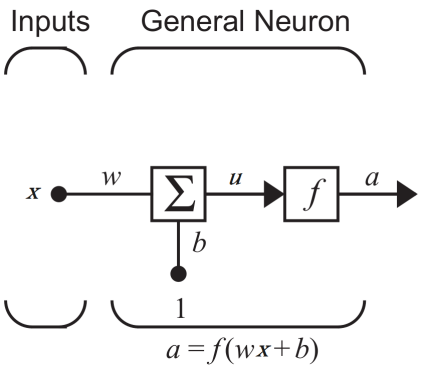

Figure 4.11: Single-input neuron. Image taken from. Image taken from [62].

Neural Network Architecture

In most of cases, the input is not a scalar but a vector (multivariate data). A connexion of multiple neurons is used to treat the input data and forms a layer of neurons as shown in Fig. 4.12. The weights are represented by the matrix $\boldsymbol{W}$

$$\boldsymbol{W} = \begin{pmatrix} w_{11} & w_{12} & \cdots & w_{1R} \\ w_{21} & w_{22} & \cdots & w_{2R} \\ \vdots & \vdots & \ddots & \vdots \\ w_{S1} & w_{S2} & \cdots & w_{SR} \end{pmatrix}$$

and the biases by a vector **b**. Each weight w_{ij} represents the connection to the neuron number $i = \{1, \ldots, S\}$ from the input number $j = \{1, \ldots, R\}$. The NN corresponds then to a set of connected layers of neurons. The layer connected to the data is called input layer, the one related to the output is known as the output layer, the remaining layers are called hidden layers. Usually, three or four layers are used to get a practical neural network. Increasing the number of the network' layer lead to a deep NN. For a network to be able to generalize, it should have fewer parameters than there are data points in the training set [62]. Fig. 4.13 shows an example of an arbitrary three-layer network.

The success of NNs can be justified by the theorem of Cybenko (1989) stating that a NN with one single hidden layer is a *universal approximator* that can represent any continuous function on compact subsets of $\mathbb{R}^n$. CNNs, recurrent neural networks (RNN), long-short-term-memory (LSTM), generative adversarial networks (GAN) and autoencoders are examples of NN architectures [10] [63] [64].

Backpropagation

The intelligence of the artificial neural network and its ability to learn consist in the adjustment of the weights between neurons so that the output can fit a desired target. The backpropagation (BP) is a common learning method used for training NNs. A set of input and the corresponding desired output (target) must be provided, so that the adjustment can be made after comparing the output to the known true target. BP consists in two steps: forward and backward.

Forward propagation: Initially, small random values are assigned to the weights. An input is then applied to the network. The input of each layer is weighted and summed (along with the bias). The resulting signal (**u**) is then passed through the activation function and transmitted to the next layer. This process is called forward propagation

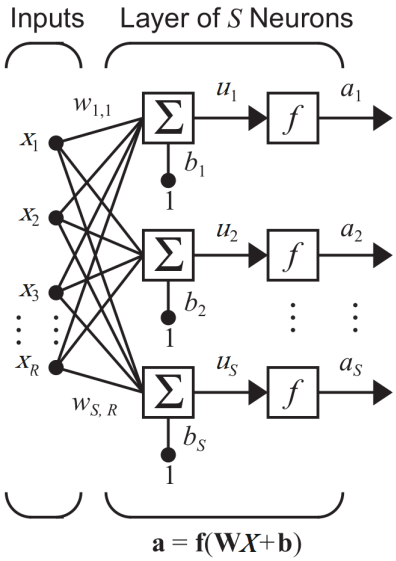

Figure 4.12: A layer of neurons. Image taken from [62].

and can be described as follows:

$$\begin{cases} \mathbf{a}^0 = \mathbf{x} \\ \mathbf{a}^{c+1} = \mathbf{f}^{c+1}(\boldsymbol{W}^{c+1}\mathbf{a}^c + \mathbf{b}^{c+1}) \quad c = 0, 1, \ldots, C-1, \\ \hat{\mathbf{y}} = \mathbf{a} = \mathbf{a}^C, \end{cases} \tag{4.2}$$

where C refers to the number of layers in the model.

Backward propagation: The process of backward propagation is a generalization of the least mean squares algorithm (LMS) which is an adaptive filter that aims to estimate the system coefficients minimizing the mean square error between the system output ($\hat{\mathbf{y}}$) and the desired one ($\mathbf{y}$) [65]. Given the output of the network and the correct target, known from the ground truth data, a performance index (H) is calculated as follows:

$$H(\mathbf{w}, \mathbf{b}) = E((\mathbf{y} - \hat{\mathbf{y}})^2) = E((\mathbf{y} - \mathbf{a})^2), \tag{4.3}$$

Using the Widrow-Hoff learning rule, Eq. (4.3) can be approximated by the squared error at iteration k:

$$H(\mathbf{w}, \mathbf{b}) \approx \hat{H}(\mathbf{w}, \mathbf{b}) = (\mathbf{y}(k) - \mathbf{a}(k))^2.$$

According to the LMS algorithm, the stochastic gradient descent (SGD) for minimizing the squared error at layer number c is given by:

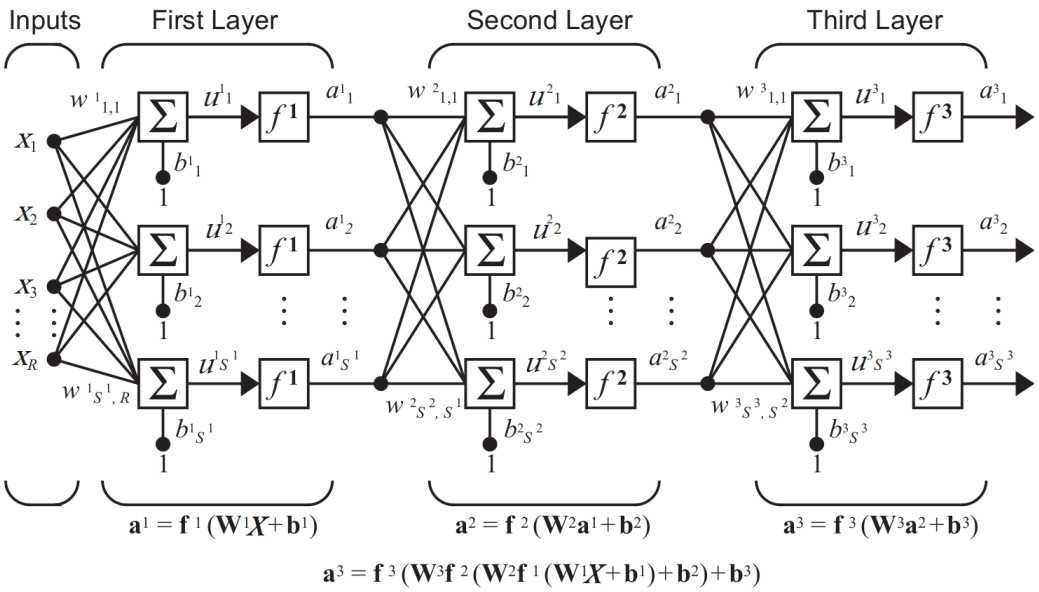

Figure 4.13: A three-layer NN. Image taken from [62].

$$w_{ij}^c(k+1) = w_{ij}^c(k) - \epsilon \frac{\partial \hat{H}}{\partial w_{ij}^c}, \tag{4.4}$$

$$b_i^c(k+1) = b_i^c(k) - \epsilon \frac{\partial \hat{H}}{\partial b_i^c}, \tag{4.5}$$

where ϵ represents a small positive real number, known as learning rate. Developing the derivatives in equations Eq. (4.4) and Eq. (4.5) with the use of the chain rule results in:

$$\frac{\partial \hat{H}}{\partial w_{ij}^c} = \frac{\partial \hat{H}}{\partial u_i^c} \frac{\partial u_i^c}{\partial w_{ij}^c},$$

$$\frac{\partial \hat{H}}{\partial b_i^c} = \frac{\partial \hat{H}}{\partial u_i^c} \frac{\partial u_i^c}{\partial b_i^c}.$$

The input vector $\mathbf{u}^c$ of the layer c can be formulated as the following:

$$\mathbf{u}_i^c = \sum_{j=1}^{S^{c-1}} w_{ij}^c a_j^{c-1} + b_i^c. \tag{4.6}$$

Deriving Eq. 4.6 results in:

$$\frac{\partial u_i^c}{\partial w_{ij}^c} = a_j^{c-1}, \frac{\partial u_i^c}{\partial b_i^c} = 1.$$

Let $s_i^c = \frac{\partial \hat{H}}{\partial u_i^c}$ represents the sensitivity of $\hat{H}$ to changes with respect to the ith input of at the c^{th} layer. As a result, the approximate steepest descent algorithm can be formulated as:

$$w_{ij}^c(k+1) = w_{ij}^c(k) - \epsilon s_i^c a_j^{c-1},$$

$$b_i^c(k+1) = b_i^c(k) - \epsilon s_i^c.$$

The matrix formulation becomes:

$$\mathbf{W}^c(k+1) = \mathbf{W}^c(k) - \epsilon \mathbf{s}^c (a^{c-1})^T,$$

$$\mathbf{b}^c(k+1) = \mathbf{b}^c(k) - \epsilon \mathbf{s}^c,$$

such that

$$\mathbf{s}^c = \frac{\partial \hat{H}}{\partial \mathbf{u}^c} = \begin{pmatrix} \frac{\partial \hat{H}}{\partial u_1^c} \\ \frac{\partial \hat{H}}{\partial u_2^c} \\ \vdots \\ \frac{\partial \hat{H}}{\partial u_{S^c}^c} \end{pmatrix}.$$

The sensitivities $\mathbf{s}^c$ are calculated by means of the following Jacobian matrix :

$$\frac{\partial \mathbf{u}^{c+1}}{\partial \mathbf{u}^c} = \begin{pmatrix} \frac{\partial u_1^{c+1}}{\partial u_1^c} & \cdots & \frac{\partial u_1^{c+1}}{\partial u_{S^c}^c} \\ \vdots & & \vdots \\ \frac{\partial u_{S^{c+1}}^{c+1}}{\partial u_1^c} & \cdots & \frac{\partial u_{S^{c+1}}^{c+1}}{\partial u_{S^c}^c} \end{pmatrix}$$

$$\begin{aligned} \frac{\partial u_i^{c+1}}{\partial u_j^c} &= \frac{\partial u_i^{c+1}}{\partial a_j^c} \frac{\partial a_j^c}{\partial u_j^c}, \forall i, j \\ &= w_{ij}^{c+1} \frac{\partial a_j^c}{\partial u_j^c} \\ &= w_{ij}^{c+1} \frac{\partial f^c(u_j^c)}{\partial u_j^c} \\ &= w_{ij}^{c+1} \dot{f}^c(u_j^c). \end{aligned}$$

Therefore

$$\frac{\partial \mathbf{u}^{c+1}}{\partial \mathbf{u}^c} = \mathbf{W}^{c+1} \dot{\mathbf{F}}^c(\mathbf{u}^c),$$

where $\dot{\mathbf{F}}^c(\mathbf{u}^c) = \mathrm{diag}(\dot{f}^c(u_1^c), \ldots, \dot{f}^c(u_{S^c}^c))$.

The sensitivity at the layer c is then computed from the sensitivity at the layer $c+1$, which explains the use of the term BP. A simplified formula is given by:

$$\begin{aligned}\mathbf{s}^c &= \frac{\partial \hat{H}}{\partial \mathbf{u}^c} \\ &= \left(\frac{\partial \mathbf{u}^{c+1}}{\partial \mathbf{u}^c}\right)^T \frac{\partial \hat{H}}{\partial \mathbf{u}^{c+1}} \\ &= \dot{\mathbf{F}}^c(\mathbf{u}^c)\left(\mathbf{W}^{c+1}\right)^T \mathbf{s}^{c+1}.\end{aligned}$$

Training a Neural Network

As explained above, a cost function H is minimized during the learning phase, Function H decomposes in a sum of per-example training losses (L) over the training samples. Let $\mathbf{y}$ contain the labels attributed to a sample, $q(\mathbf{y})$ be the unknown true distribution of the m data samples and let $p(\mathbf{y})$ be the distribution estimated by the classifier. The average training loss can be computed as:

$$H_p(q) = \frac{1}{m}\sum_{i=1}^{m} L(\mathbf{x}_i, y_i, q).$$

The cross-entropy loss function is widely used in ML and is preferred to MSE for example [10]. Binary cross-entropy is given by the following equation:

$$H_p(q) = -\frac{1}{m}\sum_{i=1}^{m}\left[y_i \log(p(y_i)) + (1-y_i)\log(1-p(y_i))\right]. \tag{4.7}$$

In information theory, the entropy measures the uncertainty associated with a given variable, $q(\mathbf{y})$ distribution here. And minimizing the cross-entropy loss aims at maximizing the fit between $p(\mathbf{y})$ and $q(\mathbf{y})$ distributions in terms of the entropy measure.

BP consists in adjusting the weights of a NN along the training process with the use of gradient descent (GD). SGD represents a stochastic extension of GD that computes H with respect to each data samples separately and sequentially one after the other. Another training approach consists in feeding the NN not only with a single input data per iteration but by a mini-batch (subset) of samples at once. This solution is in-between SGD, that might be noisy and extremely slow when dealing with large datasets, and batch gradient descent, that takes all the dataset at once and that may turn out unmanageable with huge datasets as well. Let $\mathcal{B} = \{\mathbf{bx}_1, \ldots, \mathbf{bx}_{m'}\}$ represent the mini batches drawn uniformly from the training dataset, then the total loss H is computed as the

average of the m' sub-losses:

$$H_p(q) = \frac{1}{m'} \sum_{i=1}^{m'} L(\mathbf{bx}_i, y_i, q).$$

One epoch consists in a full training iteration of GD where the NN has seen the totality of the data samples. The main limitation of the classical training with GD consists in the high sensitivity to some directions in the parameter space (with respect to ϵ). Alternative algorithms with adapted learning rates arose such as adaptive gradient algorithm (AdaGrad), root mean square propagation (RMSProp) and adapted moments (Adam) [10] [66]. Adam allows us to choose hyperparameters in a robust way, as it combines RMSProp algorithm with momentum.

The training process can be stopped in several ways:

1. Predefining a prior number of epochs after which the training finishes. However, the choice of this number is subjective and can hamper the classification performance as it can stop the training before convergence.

2. Early stopping criteria consisting in stopping the training when a specific criteria (predefined) does not evolve for at least a specific number of epochs, called patience. The stopping criteria can be for example the AUC or loss.

Basic Models

As stated earlier in this section, there exists a wide variety of NN models: CNN, RNN, LSTM, GAN and autoencoders to name a few [10] [63] [64]. The choice of the models depends mainly on two factors, the type of data format and the ML task:

- Data format: time series, image, text and multimodal input to name a few.

- Task: ECG classification, customer clustering, churn prediction and weather forecast to name a few.

A RNN is adapted to processing sequences and temporal series of data points, such as video and speech, thanks to their ability to keep memory with an internal loop. It processes sequences by iterating through the input sequence elements and maintaining a *state* containing information relative to what it has seen so far, as illustrated in Fig. 4.14.

An LSTM model is a special case of RNNs, developed in 1997, to prevent the gradually vanishing gradient during processing and to be able to learn long-term dependencies [63]. Similarly to RNNs, LSTM networks have a chain-like structure, but the

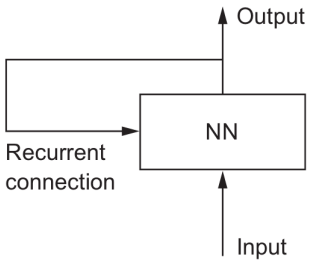

Figure 4.14: A recurrent netwok: a neural network with a loop. Image taken from [10].

repeating module has a different structure. Instead of having one NN layer, it contains four layers interacting in a particular way.

A GAN model is unsupervised and it learns to generate new data with the same statistics as the training set. For example a GAN trained on a dataset of photographies of celebrities, generates synthetic but "almost realistic" images of "fake celebrities". A GAN NN network consists in two networks: generative and discriminative. The generative network synthesizes new samples by learning how to map from a latent space to a data distribution of interest while the second network evaluates the synthetic candidates in terms of data distributions to distinguish them from the real ones. The goal behind training the generative network is to increase the error rate of the discriminative network, thus producing new samples that the discriminator network thinks are not synthesized [10] [67].

The basics of the CNN model will be detailed in Sec. 4.4.6.

Dropout Regularization

A regularization technique is used to reduce the error by fitting a function (learned by the NN during the training phase) appropriately on the given training set and to avoid overfitting. Among the existing regularization techniques, dropout [68] remains powerful and computationally inexpensive. This technique consists in dropping out a random set of nodes of the NN during each epoch, then training the remaining sub-NN. It allows one to alleviate the training computational cost and to avoid overfitting by introducing noise and forcing the non-dropped nodes to take on more or less important weights with respect to the inputs. Dropout requires the definition of a hyperparameter called dropout rate/probability, which corresponds to the ratio of the dropped vs. the total number of nodes. The dropout rate is often chosen between 0.5 and 0.8 in hidden layers and gets higher (0.8) in input layers.

4.4.6 Convolutional Neural Network

A regular NN is formed by FC layers where all neurons of a hidden layer are connected to the totality of neurons in the previous one. However, the regular NN architecture does not scale well to large and full images such as in CIFAR-10, where images are RGB and have (32×32×3) dimensions. A classical densely connected neuron in the input layer would have 32×32×3 = 3072 trainable parameters. The number of weights may become extremely high with larger images. For example, an image of size 200x200x3 would result into a number of weights equal to 200×200×3 = 120,000. It becomes clear that the dense connectivity of neurons entails a huge training load. Instead of being fully connected, the nodes in a CNN are connected only to a small region, also called patches, of the input. The dimensions of the full (input) image are reduced through the network layers into a single vector containing the scores of the different classes. The scores are arranged along the depth dimension, also called channels, and the convolutional filter is three-dimensional, as shown in Fig. 4.15. CNN can also be adapted to 1D and 2D input data. A CNN is a special NN composed of convolutional (conv) and pooling (pool) layers followed by FC ones [10] [69]. Each node of the conv layer receives some inputs, performs a dot product and optionally follows it with a non-linearity. The output of each conv neuron is called feature map. A CNN arranges the weights of its nodes in three dimensions (width, height, depth). In pooling layers, the resolution of feature maps is reduced to increase the invariance of features to distortions. CNN captures automatically salient patterns from input data to make the classification decision. During the last decade, CNNs have been intensively used in biomedical data analysis [70], mainly for ECG classification tasks [71] [72].

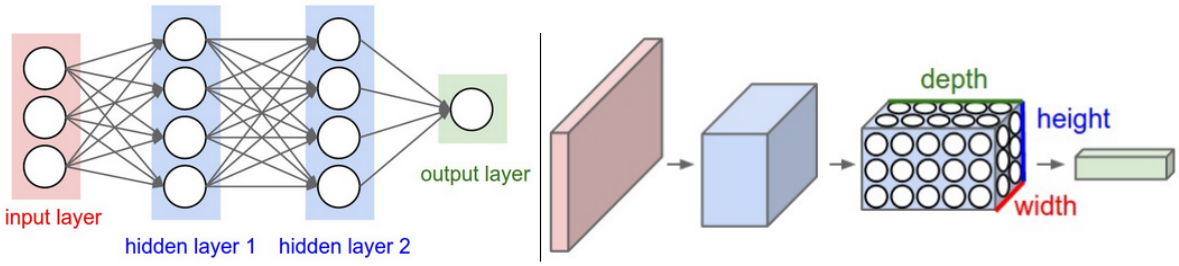

Figure 4.15: Comparison between a FC network and a CNN. **Left**: A FC 3-layer NN. **Right**: Layers of CNN transform the input volume through a 3D filter of neuron activations. In this example, the red layer represents the RGB image.[7].

[7] https://cs231n.github.io/convolutional-networks/

Convolution Operation

The parameters of conv layers consist in the learnable filters. Each filter (f) extends through the entire depth of the input patch but has small spatial dimensions (width, height). The filter size along the spatial dimensions is called *receptive field* (f_{size}). Let us suppose $f_{\text{size}} = (3 \times 3)$, then the filter is a 3D volume of dimensions $(3 \times 3 \times d)$, where d represents the depth of the input. Each filter is slided (convolved) across the spatial dimensions of the input volume. Then a dot product is computed between the weights of the filter and the elements of the patch, where the filter is positioned at that specific iteration of the forward pass. The result of convolving each filter at every spatial position of the input volume produces a 2D activation map, also called feature map. The convolution process is illustrated in Fig. 4.16. Several filters (Z) are commonly used per conv layer. As a result, the global output of the conv layer is a 3D volume consisting of the different 2D features maps, stacked along the depth dimension. The local connectivity between the conv filters and inputs allows the filters to learn activating when they see visual features such as blotches of colors, edges, orientations and even more elaborated features, such as noise and eyes in face recognition tasks, on higher (deeper) layers of the CNN. Learning to activate means developing weights of important values (non-vanishing) at the end of BP training process.

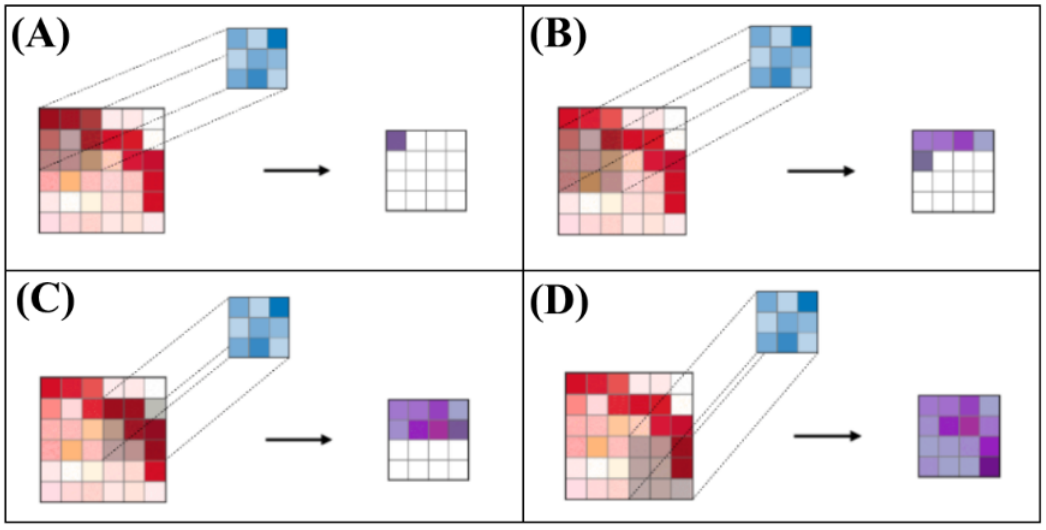

Figure 4.16: Illustration of convolutional filtering. The filter in **blue** is first positioned at the beginning of the input image in **red** (**A**), then evolves/scrolls throughout the image (**B**, **C**) until it reaches the totality (**D**). The resulting feature map is shown in **purple** and has a smaller dimension than the input (Figure modified from Standford.edu cheatsheets on convolutional neural networks).[8]

[8]https://stanford.edu/ shervine/teaching/cs-230/cheatsheet-convolutional-neural-networks

The difference between dense and conv layers remains in that FC layers learn *global* patterns from the input space, whereas conv layers break the input into *local* patterns and learns local features such as edges and textures in case of images [43]. A notion of local patterns is based on the assumption that closer pixels are more likely to be correlated. The two fundamental characteristics of convolution are translation invariance and spatial hierarchy of patterns:

- Similarly to the visual world that is translation invariant, the patterns learned by a CNN are invariant to translation. Indeed, a pattern captured in the upper-left corner of an image can be recognized by the conv network anywhere else (in the image). However, if the same pattern is shifted at a new location inside the image, a dense network would have to re-learn it anew. Hence, CNNs are data efficient as they need fewer training data samples to learn generalizable representations. Hence, they show improved generalization power.
- The successive conv layers can learn spatial hierarchies of patterns as illustrated by Fig. 4.17. Small local patterns such as edges can be learned by the first conv layers. The next layers will learn larger and more elaborated patterns made of the previously captured patterns, and so on. Hence, CNNs are able to learn increasingly complex visual concepts.

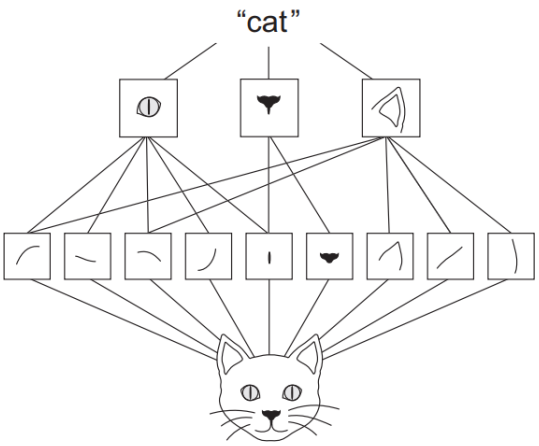

Figure 4.17: Spatial hierarchy learned by a CNN. Hyperlocal edges leaned in the first layer are combined into local objects such as ears and eyes. Combining these features yields higher-level concepts, such as a cat face. Image taken from [43].

The linear transformation (convolution) may be followed or not by a non-linear transformations such as sigmoid, rectified linear unit (ReLU) or hyperbolic tangent. Introducing non-linearity in the network allows us to model the response variable (target variable,

class label, or score) that varies non-linearly with its explanatory variables. The ReLU function [74] [75] gives the positive part of its argument (arg):

$$\text{ReLU}(arg) = arg^{+} = \max(0, arg)$$

and is a widely used.

Pooling Operation

Pooling allows to downsample data by reducing progressively the spatial size of the representation, thus reducing the amount of parameters and the computational load in the CNN. Hence, pooling allows to handle overfitting. The pool layer operates on each feature map of the input independently and resizes it spatially using a pool function. The second advantage of pooling remains the *spatial-filter hierarchies*, it allows conv layers to look at increasingly large patches (in terms of the fraction of the input size to the receptive field) [43]. Pooling operation consists in extracting sequentially windows (patches) from the input then applying a pooling function to the elements of each patch. Pooling is conceptually similar to the conv operation but it uses a hardcoded transformation instead of a learnable filter, in the case of convolution.

A common pool window size is (2×2) with a stride of 2, thus downsampling every activation map by a factor of 2. The stride refers to the number of spatial samples (time stamp in case of time series, and pixels for images) by which the conv or pool filters are slided between successive patches. Two pool functions are commonly used: maximum (max) and average (avg). They allow to summarize the most activated presence of a pattern and the average presence of a pattern, respectively. A mixture of max and average pool functions is also possible [76]. Another pooling type is proposed in the literature, called wavelet pooling [77]. Subsampling is a common approach in signal processing, it creates a reduced representation of data that still contains the most important structural features, while discarding the fine details that might be relatively useless to the task. It practice, a pool layer is commonly inserted periodically between successive conv layers in the CNN architecture.

CNN Architectures

For a given classification task, the choice of the best network architecture and hyperparameters such as the depth of the network, number of conv layers and blocks, types of activation functions and number of nodes in each layer represents a challenging problem and an active area of research. Designing an optimal architecture is generally performed empirically even though arising studies propose algorithms to estimate the

most optimal hyperparameters such as the firefly algorithm [78]. Therefore, the choice of the architecture depends strongly on the task, the amount of available data and the complexity of the information to be treated. The design may also be constrained by the computational resources because training deep and large CNN models requires a huge volume of disk space and very efficient graphics processing units (GPU). Several CNN architectures have been proposed in the literature for different tasks and purposes. The architectures used in the framework of the present doctoral thesis include: LeNet5, VGG16 and convolutional arithmetic circuits (ConvAC), and are detailed next:

(a) **LeNet5** is a baseline CNN that is both a good classifier and a computationally-affordable algorithm [79]. It has 60,000 parameters compared to AlexNet [80] for instance that has 60 million parameters [10]. LeNet5 model was proposed for the classification of the Modified National Institute of Standards and Technology (MNIST) dataset and achieved a low error of only 0.8% on the test set. MNIST includes 60,000 small grayscale images of handwritten single digits between 0 and 9 of dimensions (28×28). The architecture of LeNet5 is schematized in Fig. 4.18. It is formed by a block of two conv layers alternated by two pool layers, then two FC layers followed by one decision layer with 10 nodes (10 classes of digits).

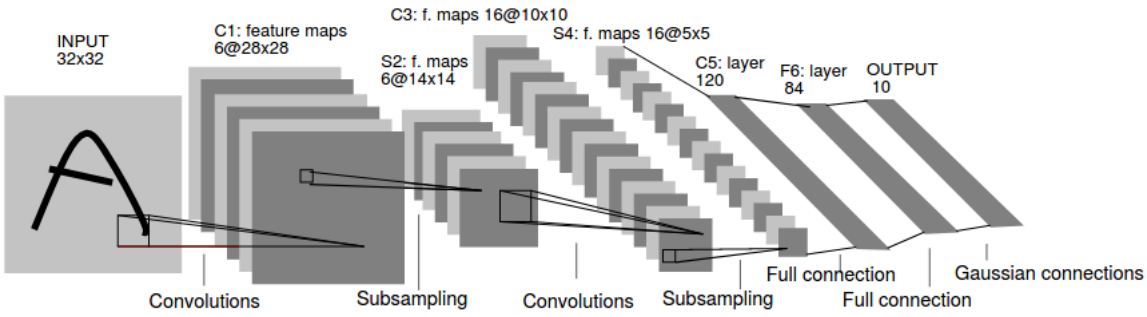

Figure 4.18: LeNet5 architecture. The input size of each layer is presented as depth @ height $\times$ width. Image taken from [79].

(b) **VGG16** is a deep CNN designed by the Visual Geometry Group Lab (University of Oxford, Oxfordshire, England), trained for ImageNet Large Scale Visual Recognition Challenge (ILSVRC) and achieving 92.7% top-5 test accuracy [81]. ImageNet is a large-scale diverse dataset containing 14 million images that belong to 1000 classes [82]. Images are natural, colored (RGB) and may have different spatial dimensions (height $\times$ width). VGG16 is composed of 5 conv blocks alternated by pooling layers and followed by FC ones as shown in Fig. 4.19. Conv blocks contain two or three consecutive conv layers with the same number of filters (within the same block), equal to 64, 128, 256, 512 and 512, respectively for the five blocks. In total VGG16 has 16 layers with trainable

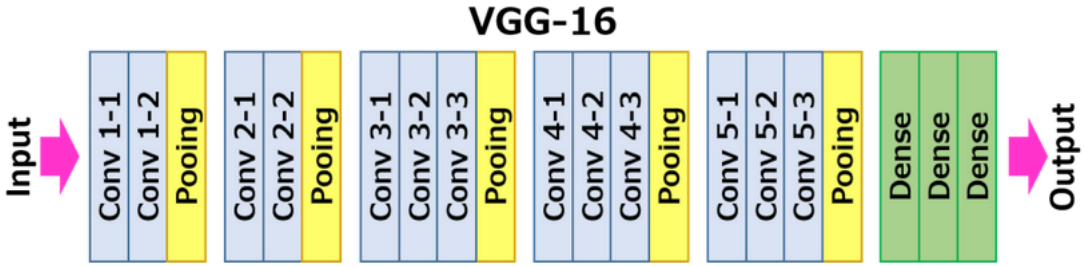

Figure 4.19: VGG16 architecture. Conv $B \times L$ represents the L^{th} conv layer of the B^{th} conv block.[9]

weights. The receptive field of the first conv layer equals (224×224) in VGG16.

(c) **ConvAC** Arithmetic circuits are Sum-Product networks [83] with two types of nodes: sum and product. A ConvAC is a CNN with:

1. Sum nodes, for linear point-wise activations implementing convolution.

2. Product nodes, for product pooling [84][85].

The interest in ConvACs is based on the theoretical formulation of their expressive power in terms of efficiency and generalization power. The originality of these networks consists in their close link with tensor decompositions. Indeed, a shallow ConvAC is equivalent to a canonical polyadic decomposition (CPD) [86] and a deep ConvAC is equivalent to hierarchical Tucker decomposition (HTD) [87]. More information about tensor decomposition will be provided in Sec. 4.9. Based on this breakthrough finding, we are interested in estimating the parameters of the ConvAC by exploiting the algebraic solution of its tensor decomposition, as explained in the next part of the thesis.
The ConvAC architecture is graphically described in Fig. 4.20. Let $\mathbf{X}$ be an instance image divided into N patches $\{(\tilde{x}_i)\}_{1,\dots N}$, a set of M representation functions $f_{\theta_1} \dots f_{\theta_M} : \mathbb{R}^s \to \mathbb{R}$ are applied to the patches, where s represents the number of pixels (elements) in a patch. The set $\{f_{\theta_d}\}_{d\in[M]}$ should be total and linearly independent in $L^2(\mathbb{R}^s)$. The transformed image is then passed to a conv layer with linear filters of receptive field (1×1), performing the scalar product with the representation function outputs, after which a product pooling is applied to the Z feature maps. A final dense layer outputs the score function h_k for $k = 1, \dots, K$, where K is the number of classes. Class k^* is attributed to $\mathbf{x}$ via maximization of per-label score function h_k.

4.4.7 Transfer Learning

In terms of ML, TL is a process where a NN (M_1) is trained on a first problem then re-used in some way in a related problem [10] [88] [89] [90]. A new model (M_2) is

[9] https://neurohive.io/en/popular-networks/vgg16/

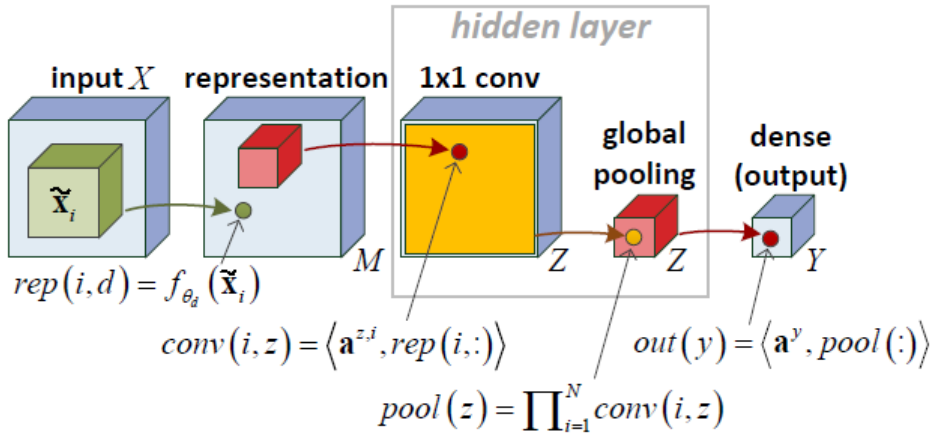

Figure 4.20: Shallow Convolutional Arithmetic Circuit. Image taken from [84].

then formed by aggregating one or several layers from the trained model M_1 to some additional layers, generally FC with trainable weights. The main benefits of TL consist in decreasing the training time for the new model and achieving a lower generalization error. TL can be used for two main purposes: feature extraction and fine tuning.

- **Feature extraction**: Only the final decision layers of M_2 have trainable parameters whereas the weights of the first reused layers, called $M_{1\text{Layers}}$, are frozen during the training process and are equal to those of M_1. As a result, $M_{1\text{Layers}}$ can be assimilated to a feature extractor. The features are those learned from the original dataset on the original task (M_1 model).
- **Fine tuning**: The weights of M_2 are initialized with those of M_1 while the weights of the final decision layers might be initialized randomly. The idea is to take profit from the features learned by M_1 as a starting point in the learning process of M_2, thus allowing the new model to learn by itself the decision rule adapted to the new dataset and the new task (M_2 model).

4.5 Performance Metrics

In order to evaluate the model's classification performance, several metrics are assessed. The accuracy (Acc) metric is computed as the result of dividing the total number of correct classifications by the total number of classifications. However accuracy is not enough to quantify classification performance when learning from imbalanced datasets. For this reason, it remains important to compute the confusion matrix whose elements are the true positive (TP), true negative (TN), false positive (FP) and false negative (FN) rates. The following metrics are evaluated:

- Sensitivity, recall or true positive rate (TPR) measures the proportion of actual positives that are correctly identified as such: $TPR = \frac{TP}{TP+FN}$.
- Specificity or true negative rate (TNR), measuring the proportion of actual negative samples that are correctly identified as such: $TNR = \frac{TN}{TN+FP}$.
- Precision or positive predictive value (PPV), assessing how trustable the result is when the model answers that a point belongs to that class: $PPV = \frac{TP}{TP+FP}$.
- Negative predictive value (NPV): $NPV = \frac{TN}{TN+FN}$.
- F_1-score stands for the harmonic mean of precision and sensitivity: $F_1 = 2\frac{PPV \times TPR}{PPV+TPR}$. It quantifies the balance between precision and recall values such that high recall accompanied with high precision values reflect that the positive class is perfectly handled by the model, low recall accompanied with high precision values reflect the model can't detect the positive class well but is sufficiently trustable when it does, high recall accompanied with low precision values reflect the positive class is well detected but the model also includes negative samples in it, and low recall accompanied with low precision values reflect the positive class is poorly handled by the model.
- The area under the Receiver Operating Characteristic (ROC) curve, shortly named AUC, is computed. The ROC curve is a graphical tool widely used to evaluate the performance of a binary classifier when varying the discrimination threshold [91].

4.6 Skewness and Kurtosis

In statistics, skewness (κ_3) is a measure of the asymmetry of a variable's probability distribution around its mean:

$$\kappa_3 = \frac{\mu_3}{\sigma^3},$$

where μ_3 is the third central moment and σ is the standard deviation of the input data. A $\kappa_3 = 0$ represents a perfectively symmetric distribution as the normal one. Two kinds of distributions are shown in Fig. 4.22 with positive (right-skewed) and negative (left-skewed) values. Both curves present tapering sides, called tails. Another measure of the symmetry property of a distribution was introduced in 2000 and is called distance skewness [94].

Kurtosis [95] refers to a measure of peakedness of a distribution. It quantifies whether the data are heavy-tailed (high κ_4) or light-tailed (low κ_4) relative to a normal distribu-

tion ($\kappa_4 = 0$). Let β_4 (≥ 1) stands for the kurtosis and is defined as:

$$\beta_4 = \frac{\mu_4}{\sigma^4},$$

where μ_4 is the fourth central moment. The excess kurtosis is defined as:

$$\kappa_4 = \beta_4 - 3.$$

Fig. 4.21 shows the probability density functions (PDF) for selected distributions with mean 0 and variance 1 presenting different values of kurtosis. The value of κ_4 of each distribution is given in the legend, along with the annotations of each PDF. The curve N (black curve) stands for a normal distribution with $\kappa_4 = 0$.

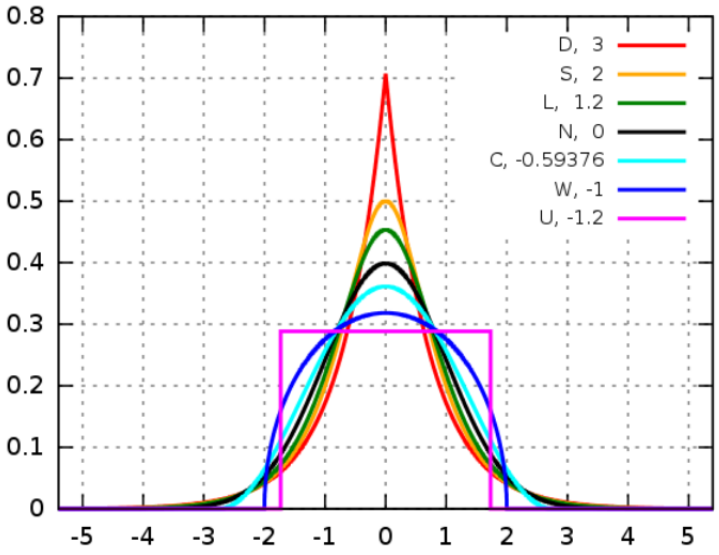

Figure 4.21: Probability density functions for selected distributions with mean 0 and variance 1 present different values of kurtosis. The kurtosis κ_4 of each distribution is given in the legend, along with the annotations of each PDF.[10]

We compute unbiased estimates of kurtosis and skewness in our experiments, detailed in the second part of the present doctoral thesis.

4.7 Cross-Validation

CV is a statistical tool commonly used in ML to quantify the generalization power of a classifier [92]. The k-fold CV consists in: first, partitioning the entire datasets into

[10] https://www.statisticshowto.com/probability-and-statistics/statistics-definitions/kurtosis-leptokurtic-platykurtic/

[11] https://codeburst.io/2-important-statistics-terms-you-need-to-know-in-data-science-skewness-and-kurtosis

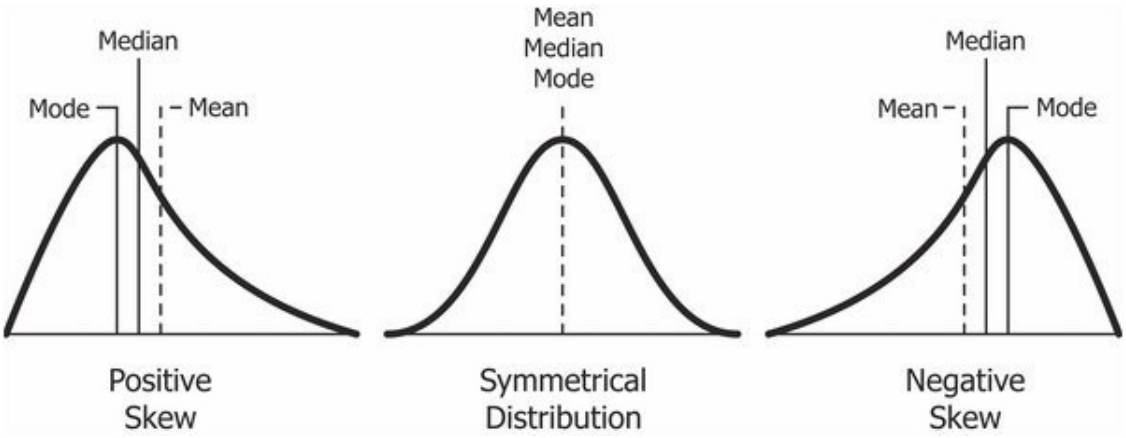

Figure 4.22: Illustration of positive, symmetrical and negative skewness distributions.[11]

k subsets called folds; second, repeating model training k times while considering, at each round, the k^{th} fold as the test dataset and the remaining samples as the training dataset; finally, the estimation of the classifier's performance is given by averaging the validation results over the k rounds. A rule of thumb is to choose k equal to 5 or 10 [93]. In each CV round, the test dataset is partitioned into two equal-sized subsets that will form the new validation and test sets. This guarantees that the model does not see the test samples during the training phase.

4.8 Data Augmentation

One of the biggest and most frequent challenges encountered in deep learning remains the insufficient amount of data or the uneven class balance of samples within the training dataset. In case of a high misrepresentation of the positive class (or negative) samples, the binary classification leads to poor results in terms of sensitivity (or specificity) and AUC. This data imbalance issue can be handled through adapted DA solutions. DA is a signal preprocessing tool that applies transformations to original samples of the minority STD class in order to synthesize new samples. It consists in forming a balanced super-dataset by replicating randomly samples from the minority class until they reach the number of the majority ones [96] [97].

4.9 Tensor Decompositions

We are interested in tensor decomposition techniques because they will be used for training ConvACs in the second part of the present doctoral thesis. The link between ConvAC transfer function and tensor decomposition will be detailed in Chap. 8.

In multilinear algebra, a tensor refers to a data structure that generalizes the representation of scalars, vectors, and matrices in higher ($\geq$ 3D) dimensional spaces. Several definitions of tensors are provided in the literature [101] [102] [103]. Motivated by our approach that is focused on ML and digital signal processing in the current doctoral thesis, we opt for the definition proposed in [103] that denotes a tensor as a multidimensional array of order G. A tensor $\mathcal{T}$ results then from the tensor product of G vector spaces such that each one has its own coordinate system. A scalar is a zero-order tensor, a vector is a first-order tensor and a matrix is a second-order tensor. A third-order tensor is a rectangular cuboid and a tensor of order $G \geq 3$ is called a high-order tensor. An arbitrary rank-1 tensor can be written as:

$$\mathcal{T} = \otimes_{i=1}^{G} \mathbf{a}^i = \mathbf{a}^1 \otimes \cdots \otimes \mathbf{a}^G, \mathcal{T} \in \mathbb{C}^{I_1, I_2, \ldots, I_G}. \tag{4.8}$$

where $\otimes$ stands for the outer product of vectors $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_G\}$. The $(d_1, d_2, \ldots, d_G)^{\text{th}}$ element of $\mathcal{T}$ is given by:

$$\mathcal{T}_{d_1 d_2 \ldots d_G} = \mathbf{a}_{d_1}^1 \mathbf{a}_{d_2}^2 \ldots \mathbf{a}_{d_G}^G.$$

A tensor has a rank Z, denoted rank(.), if it can be decomposed into the sum of a minimal number of rank-1 tensors/components, equal to Z. A tensor can be represented by matrix slices, as shown in Fig. 4.23 for an arbitrary third-order tensor. Matrix slices include all the 2D sections of $\mathcal{T}$, by fixing all but two indices of it [103].

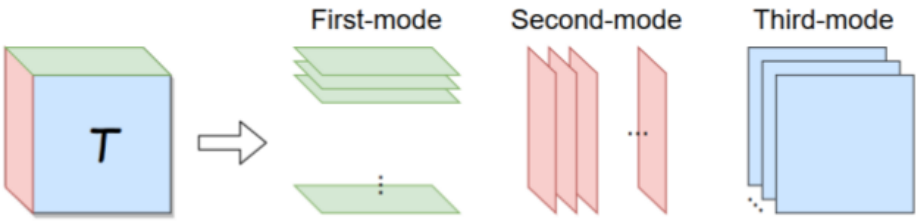

Figure 4.23: Matrix slices of a third-order tensor. Image taken from [104].

4.9.1 Canonical Polyadic Decomposition

Polyadic decomposition, also called PARAllel FACtor (PARAFAC) decomposition or CANDECOMP, is a baseline tensor decomposition form that was first introduced in 1927 [86], then developed in [99] and [100] in 1970. PARAFAC decomposes a tensor $\mathcal{T}$ into a linear combination of rank-1 tensors (outer product factors) as proposed in Eq. 4.9.

$$\mathcal{T} = \sum_{z=1}^{Z} \alpha_z \mathbf{a}^{z,1} \otimes \cdots \otimes \mathbf{a}^{z,G}, \quad \mathbf{a}^{z,i} \in \mathbb{R}^{I_i} \tag{4.9}$$

where α_z is a coefficient. The dimensions of $\mathcal{T}$ are $I_1 \times I_2 \times \cdots \times I_G$. Let $\mathbf{A}^i$ represent a matrix of size $I_i \times Z$ and $\mathbf{a}^{z,i}$ the z_{th} column of $\mathbf{A}^i$. In case the tensor $\mathcal{T}$ is decomposed into a minimal number (Z) of rank-1 tensors, PARAFAC is called *canonical polyadic decomposition* (CPD). The CP-rank refers to the minimal value of Z for which the model holds, as illustrated by Fig. 4.24, and is typically considered as the tensor rank.

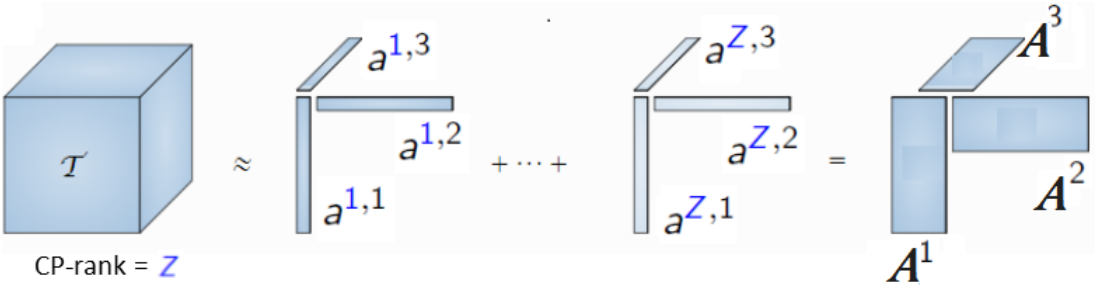

Figure 4.24: CPD of a third-order tensor.

4.9.2 Hierarchical Tucker Decomposition

The hierarchical representation of a tensor is obtained by a hierarchy of bilinear decompositions, implemented in a recursive way. Let us consider the tensor product space $\mathbf{V}$ formed by the Hilbert spaces V_j ($1 \leq j \leq G$):

$$\mathbf{V} = V_1 \otimes V_2 \otimes \cdots \otimes V_G.$$

A space is said of Hilbert if it provides scalar products $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_{V_j}$. Hence, $\mathbf{V}$ is also a Hilbert space. Hierarchical Tucker Decomposition (HTD) consists in decomposing a tensor into a new vector space obtained by a hierarchical decomposition of the input vector space into L subspaces $\mathbf{U}^{l,j}$ for $l = 1, \ldots, L$. For ease of presentation, the order G is assumed to be a power of 2 ($G = 2^L$). Let us denote $U^{0,j} \subset \mathbf{V}^{0,j} = \mathbf{V}^j (1 \leq j \leq G)$ vector spaces associated with the level $l = 0$. These subspaces form the tensor subspace $\mathbf{V}^0$ given by:

$$\mathbf{V}^0 := U^{0,1} \otimes U^{0,2} \otimes \ldots U^{0,G} \subset \mathbf{V}. \tag{4.10}$$

Based on Eq. (4.10), $\mathbf{V}^0$ can be rewritten as follows:

$$\mathbf{V}^0 = \left(U^{0,1} \otimes U^{0,2}\right) \otimes \left(U^{0,3} \otimes U^{0,4}\right) \otimes \cdots \otimes \left(U^{0,G-1} \otimes U^{0,G}\right).$$

For each pair (denoted by $V^{1,j}$), additional finitely dimensional subspaces are introduced:

$$U^{1,j} \subset V^{1,j} := U^{0,2j-1} \otimes U^{0,2j}, \quad (1 \leq j \leq 2^{L-1} = G/2).$$

This procedure is repeated for $l = 2, \ldots, L$, recursively, following Eq. 4.11.

$$U^{l,j} \subset V^{l,j} := U^{l-1,2j-1} \otimes U^{l-1,2j}, \quad (1 \leq j \leq 2^{L-l}). \tag{4.11}$$

Fig. 4.25 illustrates the hierarchical tree structure of arbitrary L vector spaces. The second step consists in calculating orthonormal bases spanning the subspaces $U^{l-1,j}$. Finally, HTD approximates the input tensor into the approximated space $\mathbf{V}$.

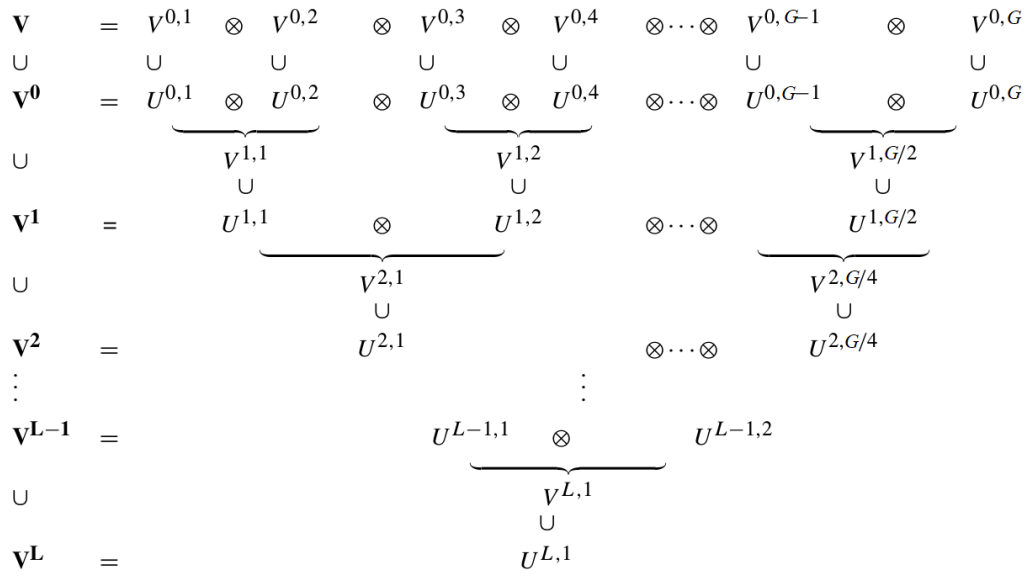

Figure 4.25: Hierarchical tree structure for $G = 2^L$. Image taken from [87].

4.10 Summary

The goal of this chapter was to introduce the mathematical tools that will be used in the rest of this thesis, in order to facilitate the understanding of the experiments detailed in Part II of the present manuscript. First, we provided the basics of clustering and classification algorithms and we focused on special CNN architectures such as LeNet5 and VGG16. The concept of TL was also introduced. Second, statistical tools to analyze the distribution of data were given such as the skewness and kurtosis. Finally, we defined CPD, a commonly used tensor de composition technique. The present chapter gave an overview of ML algorithms, statistical tools and linear algebra fundamentals that will be used for the classification of EGMs into STD vs. non-STD and the identification of potential ablation sites in STD-guided ablation. Next, the second part of the thesis will present our main contributions. It starts by introducing the study dataset that will be processed during the experiments.

Part II

Contributions

Chapter 5

Study Datasets

Contents

5.1 Introduction

The first part of this thesis gave an overview of our work and defined our main objectives. The clinical context was presented by focusing on both AF as the most clinically sustained arrhythmia and STD-guided ablation as a breakthrough ablation therapy of AF yielding promising results. The algorithms and statistical tools used to design our intelligent decision-aid solution for automatically characterizing STD EGMs were also introduced. The present chapter describes the datasets studied in the experiments that will be detailed in next chapters. The first dataset contains real EGM signals acquired from Nice Pasteur CHU and the second one is synthetic, generated in controlled conditions to asses specific aspects of the STD pattern.

5.2 Real Data

Real data are provided by interventional cardiologists from Nice Pasteur CHU, whose Ethics Committee approved the proposed research. In partnership with engineers from Biosense Webster, we were able to study real and annotated EGM data coming from persistent AF patients that are mapped with the PentaRay catheter for STD-based ablation. The data are exported from the CARTO 3 system (version 7), whose mechanisms are described in Sec. 3.2.1. The PentaRay catheter used for mapping has an edge-to-edge inter-electrode space equal to 2 mm - 6 mm - 2 mm, as shown in Fig. 3.3 and previously introduced in Sec. 3.2.3. The ablation catheter is the Thermocool SmartTouch SF 4 mm.

Studying real data is of a great interest in biomedical projects like ours as it describes better the AF specificities and provides a diversity of examples of the STD pattern. This richness of information makes the study more realistic and reliable. However, the exploitation of these real data is challenging as it requires a lot of time and a rigorous process of integrity checking, anonymization and understanding.

5.2.1 Ablation Protocol

For confidentiality reasons, we were not able to export the data of AF patients ablated with the properly defined STD-based approach in Nice CHU as part of the study [9]. It is important to highlight that ablating only atrial sites based on the STD pattern alone is still a non-conventional ablation protocol for treating AF, thus not commonly adopted in practice in Nice CHU, aside from the clinical trial conducted in [9]. As a result, the data exported belong to patients who have been ablated as follows:

1. First, ablating the triggers initiating AF by isolating the pulmonary veins.
2. Second, ablating the drivers that consist in atrial sites *of interest* (significant anatomical location) displaying STD EGMs.

This ablation combines both the principle of the classical approach (PVI) and the patient-tailored STD-guided ablation, as introduced in [9]. Hence, it remains a patient-tailored therapy. As explained in Sec. 3.3.3 and illustrated by Fig. 5.1, an ablation intervention is a two-stage process:

1. **Mapping phase**: The PentaRay catheter is sequentially positioned in various regions of the atria, then 10 EGMs are simultaneously recorded per location. Locations exhibiting a non synchronous cardiac activity are said to be STD points and are tagged for ablation. The set of all mapped points is denoted $\mathcal{P}$.
2. **CA phase**: The second step consists in ablating AF triggers and drivers that have not necessarily been mapped in advance. The set of ablated points is called $\mathcal{V}$ and

contains points from $\mathcal{P}$, but may contain also other points which are not in $\mathcal{P}$. All information about the ablated points are found in a subsection of the BioSense CARTO system, called VISITAG.

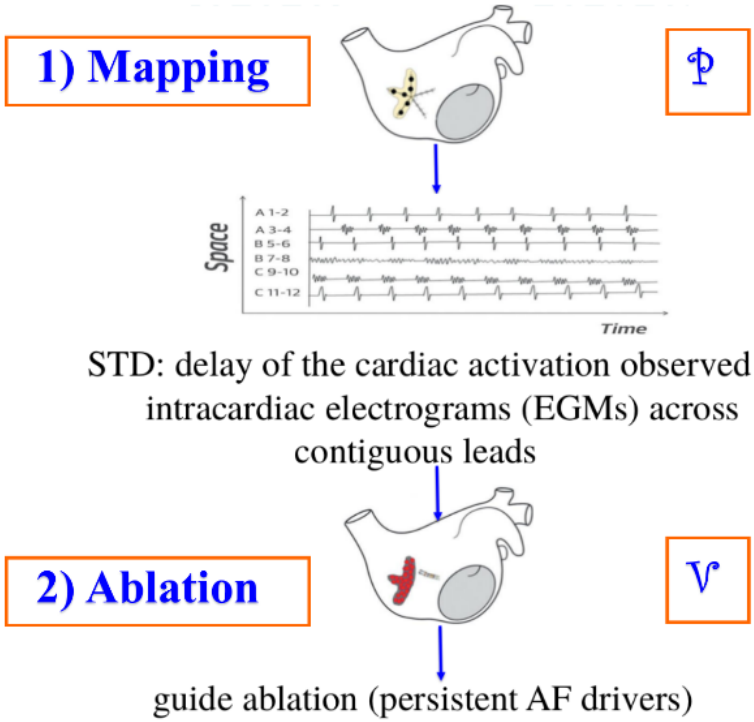

Figure 5.1: Steps of ablation intervention.

During the mapping phase, data annotation is performed by interventional cardiologists such that EGMs presenting spatiotemporal dispersion are annotated as “STD", gradient (Gra), dispersion (Dis) or substrate. Other tags are given to non-STD EGMs such as “scar", that represents a point where no electrical activity happens, generally after a previous ablation or for physiological factors. An example of multichannel EGM recordings of length 2.5 s and belonging to the STD class is illustrated in Fig. 5.2.

5.2.2 Data Exportation

The process of data acquisition consists in several steps:

1. Selecting persistent AF patients mapped with the PentaRay catheter for STD-based ablation, following the protocol described in Sec. 5.2.1. At the beginning of our research study, we identified 100 patients of interest and started by exporting the data of only 16 in order to make preliminary experiments, understand the data structure and provide a proof of concept about the feasibility of automatic classifi-

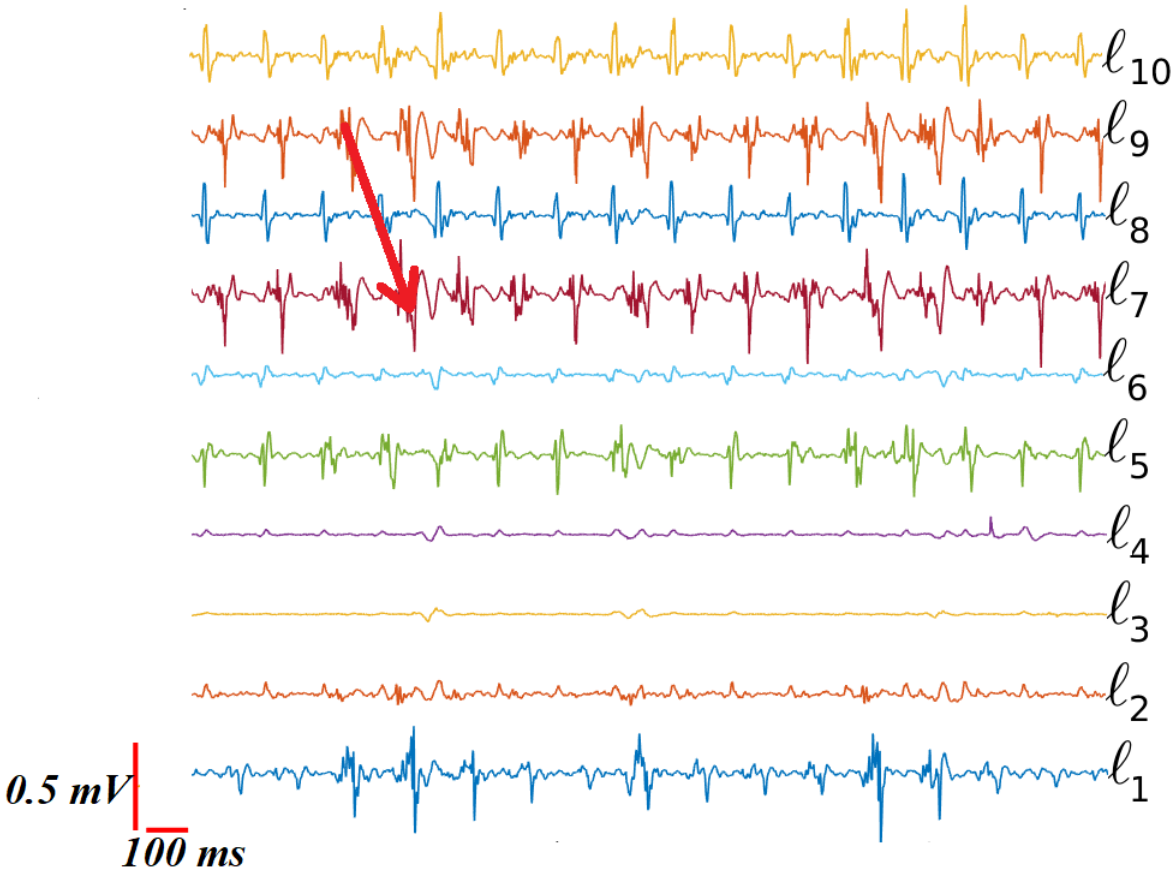

Figure 5.2: Example of 10-channel STD EGM. Contiguous leads ℓ_7, ℓ_8 and ℓ_9 display a delay of intracardiac activation, as highlighted by the red arrow.

cation of EGMs using ML and statistical analysis. However, the remaining patients identified at the beginning of the project from Nice CHU could not be exported because of the sanitary crisis related to the COVID-19 virus.

2. Exporting intervention data from BioSense CARTO system. For each patients, the data are structured as follows:

   - There are several cartographies (maps) belonging to the first or repeated mapping of the RA, LA and others.
   - Each mapping may contain hundreds or thousands of mapped locations (anatomical points).
   - Each anatomical point has Cartesian coordinates, ECGs, annotated EGMs and other types information.
   - In addition to the mapping data, we export the ablation data of each patient from the VISITAG system where we find the Cartesian coordinates of points from the set $\mathcal{V}$ added to complementary information about the contact-force used in CA with RF energy.

3. Decompressing, anonymizing and preparing data (maps/signals) of interest.

4. Structuring data for classification purposes.

Data Labeling

Data labeling is performed by interventional cardiologists as follows. EGMs presenting spatiotemporal dispersion are annotated as “STD", also called gradient or substrate. We automatically merge the remaining labels into the “non-STD" class. Moreover, a meticulous work is needed for standardizing the labels because the annotation nomenclature exported from CARTO system (encoded labels) may differ from one patient to another for technical reasons. Indeed, the dataset should have a unique label for each class (STD, scar, ...) across the different patients. In order to handle this issue, we make manual correspondence between the encoded/numerical labels (1, 9, 12, 16, 18, ...) as they are exported from the CARTO system and the tags (STD, Grad, Dis) as they are displayed on CARTO's user interface. Finally, we generate a binary label for each mapped point containing 1 if “STD" and 0 else.

Data Preparation

When exported, data are stored in different files, folders and subfolders. In order to make experiments easier and data more accessible, we propose to group each patient's

data of interest in the same structure (object) whose instances are anatomical points and attributes are:

- The 10-channel EGMs stored in a matrix.
- A rescaling coefficient.
- A table of correspondence between CARTO tags and the encoded labels.
- Binary STD label obtained after the standardization of annotation described above.

The workflow of data structuring is illustrated in Fig. 5.3. At an advanced level of ML experiments presented in the next chapters, we need to use data generators to fit in the GPU's memory. For this reason, we rethink data structuring by storing each 10-channel EGM record belonging to each patient in a separate file. The corresponding label is also stored in a separate file. Generators are widely used in DL when dealing with large and heavy datasets like the one analyzed in this study.

5.2.3 Study Dataset

The baseline information of the studied patients is given in Tab 5.1. The population is aged 64 years in average and is composed of 80% male and 20% female. The average initial AF cycle in the left atrial appendage is 156 ms. In average, the ablation intervention lasts 2 h 36 min and requires 155 RF shots that last 53 min 57 s.

Feature	**Value (mean $\pm$ std)**
Gender	12 Male + 3 Female
Age (year)	64 ± 12
Initial FA Cycle in OG (ms)	156 ± 12
Procedure Duration	2h 36min $\pm$ 42min
RF Duration	53min 57s $\pm$ 12min 15s
Number of RF shots	155 ± 61

Table 5.1: Baseline information about patient's population.

The data of 1 patient among the 16 selected were discarded in some experiments because we found incomplete VISITAG information for that patient, due to an exportation problem. Cartographies of both right and left aria are merged in the present study. The dataset includes a cohort of 35563 10-channel EGM signals of length 2.5 s composed of 1804 samples labeled STD and 33759 non-STD samples. Let CIR be the class imbalance ratio:

$$\text{CIR} = \frac{\text{\# STD}}{\text{\# non-STD}} \approx 5\%,$$

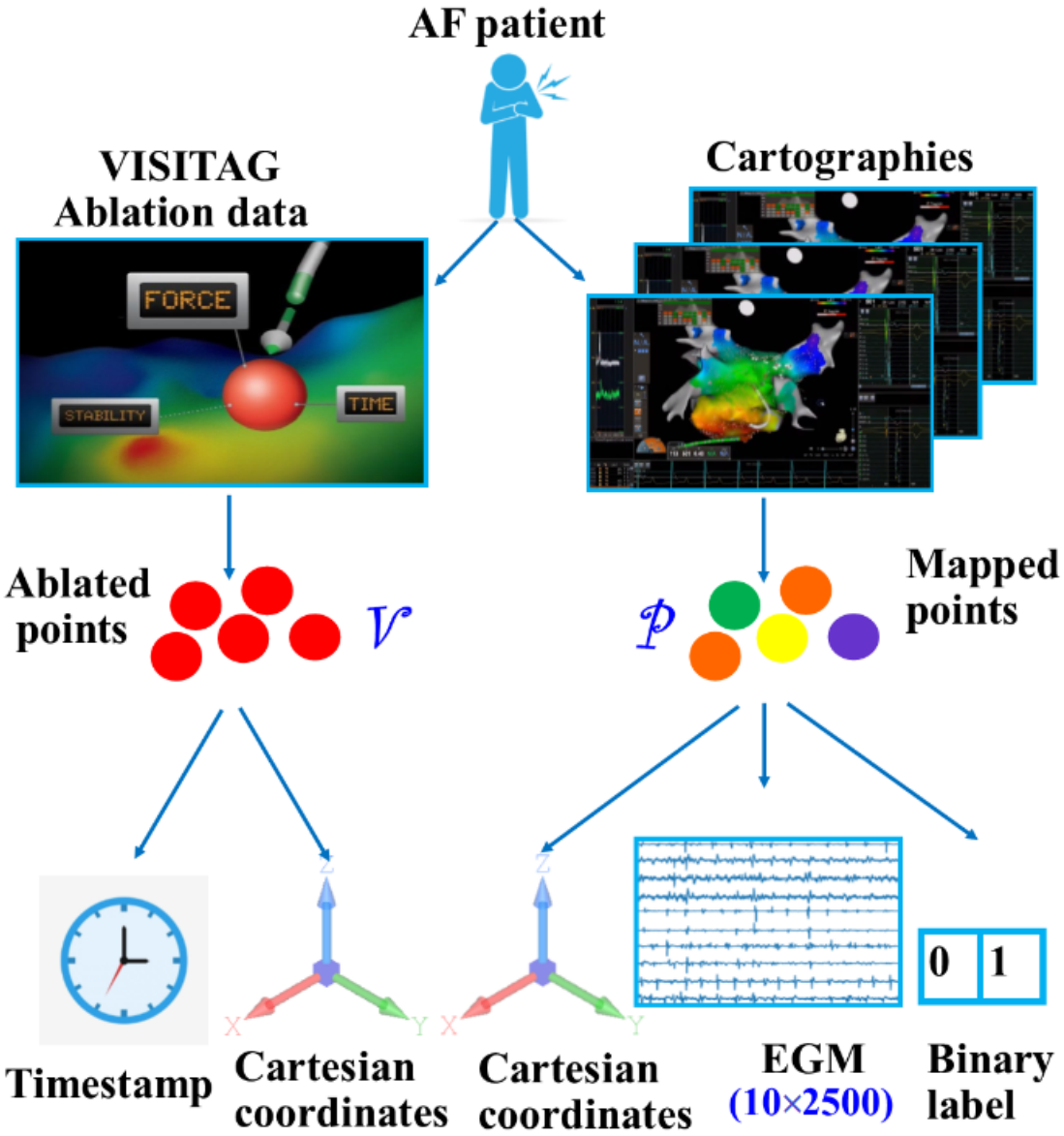

Figure 5.3: Data structuring workflow.

where # refers to the number of samples. Multichannel EGMs are sampled at 1 kHz and displayed through the CARTO system monitor, to be analyzed by the cardiologist. Hence, the exported EGM data can be stored in a matrix of dimensions 10×2500.

5.2.4 Circularity Transformation

Circularity transformation consists in replicating for each sample the first two channels (matrix rows) at the end of the sample. The new EGM matrix sample passes from dimensions of 10×2500 to 12×2500. This transformation allows us to mimic the branch circular structure (invariance) of the PentaRay catheter. Indeed, it captures the neighborhood information between bipoles (A 1-2 and A 3-4) and (E 17-18 and E 19-20), as illustrated in Fig. 5.4 and Fig. 5.5, so that bipoles in contiguous branches of the PentaRay catheter also appear in contiguous rows of the transformed matrix.

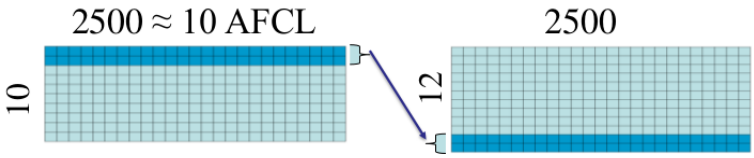

Figure 5.4: Circularity transformation.

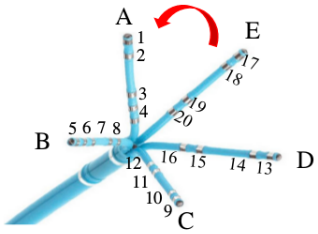

Figure 5.5: Circular structure of PentaRay branches.

5.3 Synthetic Data

To build a realistic synthetic EGM model, we simulate the cardiac activation pattern as follows:

- A 1D sawtooth activation with a fundamental frequency $F_{\text{ref}} = 50$ Hz and sampling frequency $f_s = 1000$ Hz is synthesized.
- From this full signal, a unique sawtooth cycle (one “tooth") is truncated.
- The pattern is padded with zeros to form a single cardiac cycle of 250 samples.

- This signal is replicated 10 times to form a full synthetic EGM of length 2500 noted EGM_{saw}.
- In order to include the fractionation pattern in the EGM_{saw} recording, we define d as the fractionation degree. The fractionated signal $\text{EGM}^{d}_{\text{saw}}$ should contain in each AFCL interval d squashed sawtooth waveforms, generated by truncating and padding a sawtooth waveform of fundamental frequency equal to $d \times F_{\text{ref}}$, as described previously.
- $\text{EGM}^{d}_{\text{saw}}$ is then multiplied by a normalization factor equal to $\sqrt{\frac{E_{\text{ref}}}{E}}$, thus ensuring that the energy of the time discrete signal $\text{EGM}^{d}_{\text{saw}}$ is preserved for all values of $d \in \mathbb{N}^{*}$. Let E be the energy of a time discrete signal EGM_{saw}:

$$E = \sum_{t=1}^{t=2500} |\text{EGM}_{\text{saw}}(t)|^2 = 10 \times \sum_{t=1}^{t=250} |\text{EGM}_{\text{saw}}(t)|^2.$$

The non fractionated signal corresponds to $d = 1$ and the fractionated signal corresponds to $d > 1$. Finally, the full synthetic multichannel EGM matrix is formed by stacking the synthetic signals $\text{EGM}^{d}_{\text{saw}}$ in the rows of a (10×2500) matrix. Each signal (channel) may have a different value of d as illustrated in Fig. 5.6. The 10 leads are denoted ℓ_1 to ℓ_{10}. In order to simulate the activation delay characteristic of STD, we introduce a cyclic shift of 35% of AFCL over two successive channels (matrix rows). As a result three consecutive leads would have a total delay of 70%. Fig. 5.6 shows that the delay slope may be positive (leads ℓ_3 ℓ_4, ℓ_5) or negative (leads ℓ_7, ℓ_8, ℓ_9). The model accepts different delay percentages and is not restricted to 35% between two consecutive leads.

5.4 Summary and Conclusions

The present chapter started by describing the acquisition process of the real dataset from the cardiology department of CHU of Nice. It contains a total number of more than 35000 multichannel EGM recordings annotated into STD vs. non-STD and complementary details about ablated sites and baseline information about the patients. Then, we propose a model to synthesize multichannel EGM signals containing a synthetic STD pattern, including design hyperparameters such as the number of delayed leads and fractionation degrees. Finally, the circularity transformation was introduced aiming to explicitly reflect the circular structure of the PentaRay catheter on the EGM data matrix. In the next chapter, we will tackle the experiments and the most relevant results, starting by applying data augmentation techniques to handle the high CIR (5%), as mentioned in Sec. 5.2.

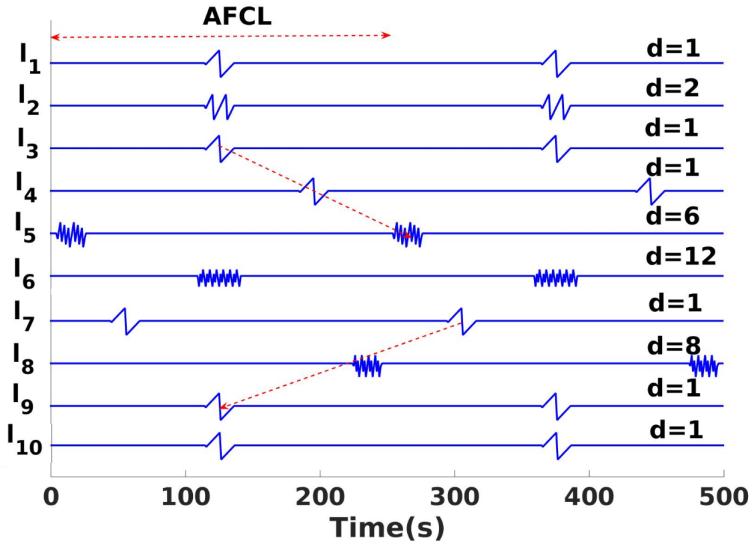

Figure 5.6: Illustrative example of a synthetic multichannel EGM recording with different fractionation degrees d.

Chapter 6

Data augmentation for Electrogram Classification

Contents

6.1 Introduction

The study datasets were introduced in the previous chapter. They contain a real dataset of more than 35000 10-channel EGM recordings annotated by interventional cardiologists into STD vs. non-STD categories. The present chapter describes the first attempt to classify a part of the real EGM samples in order to automatically identify STD EGMs using two baseline ML techniques. The main problem that we encounter in the training process is that the dataset is highly imbalanced with a CIR of 5%, as mentioned in Sec. 5.2.3. The issue of insufficient amount of training samples from the STD class is addressed through a variety of data augmentation, DA, solutions. We restrict the study to DA techniques that preserve the integrity of EGM data, mainly the STD pattern. We also present two specifically adapted DA methods called lead shift and time shift.

6.2 Classification Model

As mentioned in Sec. 5.2, for each cartography of each patient, 10 EGM bipoles are recorded per location (anatomical point inside the heart) with the PentaRay catheter. As each recorded EGM spans 2.5 s sampled at 1000 Hz, data can be stored in a matrix of dimensions 10×2500 which can be considered as an image.

As a first attempt in the development of our decision-aid solution whose aims are detailed in Sec. 1.2, we propose to classify the raw multichannel EGM recordings into STD vs. non-STD categories with the use of two baseline ML classifiers: 1) LeNet-STD, a CNN inspired by the state-of-the-art LeNet5 architecture [79]; 2) MLR. CNNs are commonly used in biomedical data classification [70]. Each node of the conv layer receives some inputs, performs a dot product and optionally follows it with a non-linearity. LeNet-STD is a CNN inspired by the previously described LeNet-5 architecture, introduced in Sec. 4.4.6, which represents a good and computationally-affordable model. Let Z and kernel respectively refer to the number of filters and the receptive field of the filters. LeNet-STD is composed of the following layers:

1. 2D conv with Z=32 and kernel=4×300.
2. Average pool.
3. 2D conv with Z=16 and kernel=3×3.
4. Average pooling.
5. FC with Z=128.
6. FC with Z=64.
7. Final FC decision with Z=2.

Logistic regression model estimates the probability of a given class by using the logistic function, as explained in Sec. 4.4.3. It represents a frequent choice in many medical data classification tasks [56], and is hence used as a reference here. The ML model is trained to recognize dispersion patterns in contiguous electrodes and apply the rule to unseen data. A natural choice for the input convolutional filter size is $4 \times \alpha_1$AFCL ($\alpha_1 \geq 1 \in \mathbb{Q}$). We opt for $\alpha_1 = \frac{3}{2}$ in order to capture the cardiac activation along one and a half AFCL. Indeed, dispersion is observable on a minimum of three electrodes positioned on two contiguous PentaRay branches and the average AFCL is typically 200 ms. For each sample, the circularity transformation is applied (Sec. 5.2.4) and results into a matrix of dimensions 12×2500.

The metrics used for evaluating the models classification performance are the following:

1. Binary cross-entropy, defined in Eq. (4.7), Sec. 4.4.5.
2. Area under the ROC curve (AUC).
3. Sensitivity (TPR), measuring the proportion of actual STDs that are correctly identified as such.
4. Specificity (TNR), measuring the proportion of actual non-STDs that are correctly identified as such.

The definition of the metrics were previously introduced in Sec. 4.5. The LeNet-STD and MLR models are trained with a mini-batch gradient descent optimizer, the mini-batch size is 32. The early stopping criterion used for training is the validation AUC for a patience of 5 epochs, as explained in Sec. 4.4.5.

6.3 Classification With Data Augmentation

The experiments described in this chapter are conducted on the EGM samples belonging to the LA cartographies alone, thus a total number of 23082 10-channel EGM recordings, composed of 1139 STD and 21943 non-STD samples. As a first attempt, we train the classification models, specified in Sec. 6.2, using the imbalanced dataset and we obtain a validation, train and test accuracy of 95% for LeNet-STD and MLR as shown in Tab. 6.1. However, accuracy alone is not enough to evaluate the classification performance, since the values of sensitivity and specificity are highly imbalanced. Indeed, the networks classify properly the majority class (non-STD) with a very high specificity equal to 98% and 99% using LeNet-STD and MLR respectively but fails to identify the minority STD samples, thus entailing a low sensitivity of 50% and 6% using LeNet-STD and MLR respectively. This problem is due to the uneven class balance and the lack of a sufficient number of STD samples needed for a fair learning process. A way of dealing with this issue is to apply DA methods that are adapted to the medical question and that are approved by cardiologists. DA [96] [97] consists in synthesizing new samples belonging to the minority class by applying transformations to its original samples, as explained in Sec. 4.8.

Classifier	**accuracy**	**sensitivity**	**specificity**	**AUC**
LeNet-STD	0.953	0.505	0.978	0.903
MLR	0.950	0.056	0.999	0.835

Table 6.1: Test classification performance for models trained on the imbalanced EGM dataset.

6.3.1 State-of-the-Art Techniques

We find in the literature several studies addressing the issue of medical data imbalance or the lack of a sufficient number of training samples with the use of DA. DA techniques can be classified into two main categories, traditional and DNN-based:

1. Traditional techniques consist in applying affine transformations or color modifications (if data samples are images). Fig. 6.1 illustrates some affine transformations such as rotation, reflection, scaling and shearing.

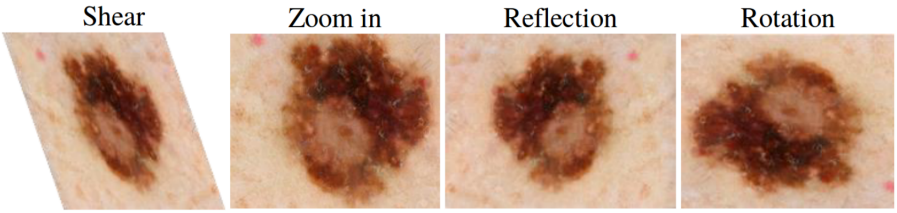

Figure 6.1: Same image after different types of affine transformations. Image taken from [96].

2. Black-box techniques are based on DNNs, mainly GAN networks. As explained in Sec. 4.4.5, GANs allow us to generate synthetic images in different ways such as:

   - Obtaining super-resolution images out of low-resolution ones.
   - Blending, obtained by mixing selected parts of two input images to form a new one.
   - Inpainting, obtained by restoring missing pieces of an image.

   A GAN model is used in [105] to augment a training dataset of computed tomography (CT) images by generating synthetic samples without contrast out of the original ones with contrast. The study aims to segment CT images of organs with DNNs and uses DA to increase the size of the dataset while introducing noise (contrast information). The results show that DA with GAN allowed to increase the performance metric from 9% to 66% in the segmentation of kidney CT images.

The issue of data imbalance in AF identification is addressed in [107]. The study aims at classifying a dataset of 8528 ECG recordings acquired from the PhysioNet challenge 2017 with a deep RNN while balancing the number of samples among different classes and increasing the dataset diversity. The goal of PhysioNet Challenge 2017 is to differentiate AF from noise, NSR or other rhythms in short term ECG recordings [108] [109]. The public dataset contains 5154, 771 and 2603 ECGs belonging respectively to the following categories NSR, AF and other rhythms. The DA algorithm used concatenates the

duplicated ECG episode to the original one based on characteristic points and resamples new recordings randomly from the concatenated signal. The average F1 score over 10-fold CV evolved from 36% to 76% with the proposed DA algorithm. DA can also be used to improve the robustness of classification models by adding perturbation to input data. For example, random combinations of DA techniques are applied to the public dataset of PhysioNet challenge 2017 and allow to improve the F1 score by 3.17% compared to the baseline model (without DA) [106]. The classification model is a deep CNN with 15 conv blocks added to residual connections. Among the DA techniques selected randomly in the study [106], we find:

- Erase: Randomly select a lead and set the signal of the selected lead to 0.
- Flip: Flip the signal up and down at random.
- Cutout: Set a random interval signal to 0.
- Sine: Add a sine wave to the entire sample.

6.3.2 Proposed EGM DA Techniques

In order to handle the low values of CIR and TPR faced in our experiments, we propose to use DA techniques that:

- Are adapted to the medical problem.
- Preserve the integrity of EGM data.
- Preserve the characteristics of the STD pattern.

The main techniques used include a set of state-of-the-art techniques added to specifically designed ones and can be summarized as follows:

1. **Undersampling**: forming a balanced sub-dataset by keeping all STD samples and selecting randomly as many non-STD samples while discarding the rest [97]. This is actually a data reduction technique.
2. **Oversampling**: forming a balanced super-dataset by replicating randomly STD samples until they reach the number of non-STD ones [97], commonly called random oversampling and denoted ROS. Similarly to undersampling, oversampling is implemented with the use of the imbalanced-learn Python library[1].

[1] https://github.com/scikit-learn-contrib/imbalanced-learn

3. **Lead shift**: recursively rotating the EGM channels so that each sample results into 9 new samples. Exploiting the circularity of the PentaRay branches, we consider the 10 channel version of EGMs (10×2500) and roll (cyclic shift) each one into 9 new versions. Finally, the circularity transformation is applied to the synthetic matrix as mentioned in Sec. 6.2. This transformation is illustrated by Fig. 6.2.

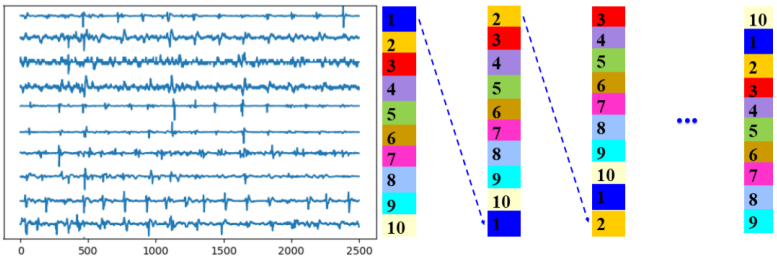

Figure 6.2: Illustration of lead shift transformation.

4. **Time reversing**: doubling the number of STD samples by flipping along the temporal dimension each sample from the end to the beginning, as if time inverting the heart beating mechanism. This operation is not expected to harm the STD identification process since the STD pattern is conserved, as if the propagation pattern traveled in the opposite direction [97]. This transformation is illustrated by Fig. 6.3.

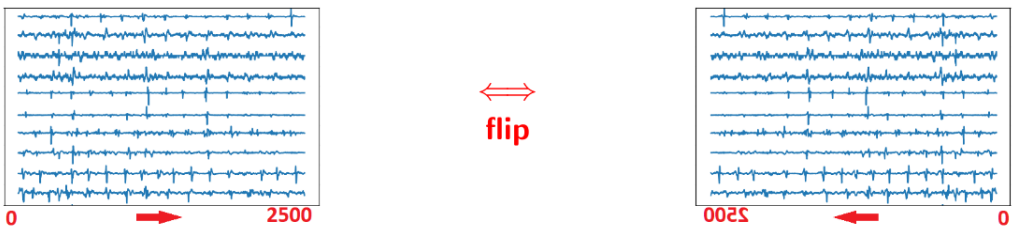

Figure 6.3: Illustration of time reversing transformation.

5. **Time shift**: truncating STD images of size 12×2500 in time dimension to form sub-images of size 12×2250. The truncation limits are shifted each time by 25 samples along the temporal dimension, that is about 10% of typical AFCL values. As a result, each original STD image results into 10 synthetic samples. This transformation is illustrated by Fig. 6.4.

The original dataset samples are split into 70% for training, 15% for validation and 15% for test sets. Then, DA is only applied to training samples. The test and validation sets should be untouched to preserve the original data distribution. In the case of time shifting, only the dimensions of the test and validation samples are reduced to 12×2250

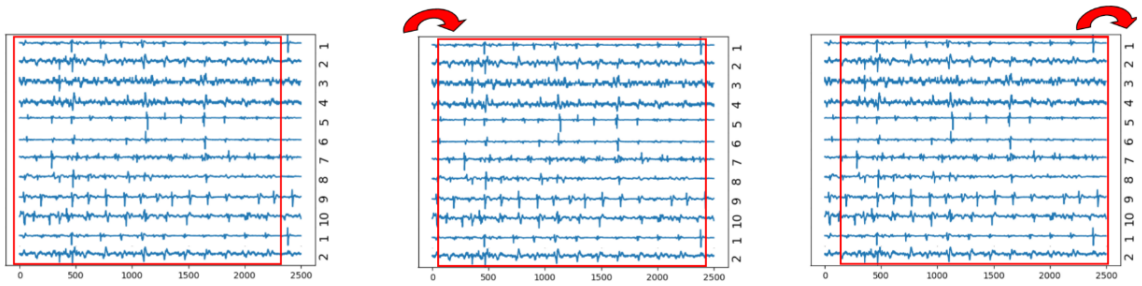

Figure 6.4: Illustration of time shift transformation.

in order to be coherent with the dimensions of the training samples.

Implementation: We conduct these ML learning experiments using Python 3.6.8 programming language, Tensorflow 2.0 deep learning library for GPU and Keras API within Docker environment, running on nVidia GeForce 930 MX GPU and Ubuntu 16.4 operating system.

6.4 Results

We benchmark the performances of LeNet-STD and MLR classifiers using different combinations of DA methods. The classifier is first trained on imbalanced dataset (no DA) then on augmented datasets. Combinations of DA techniques are also assessed as listed in Fig. 9.2. For instance, "lead shift-oversampling" refers to augmenting the training set (X) with lead shift to obtain a larger set (X^1) then augmenting X^1 with oversampling (X^2), the model is then trained on X^2. The CIR refers to the number of STD over non-STD samples (Tab. 6.2), as defined in Sec. 5.2.3. Using a DA method followed by undersampling means augmenting the number of STD samples before downsampling the non-STD class.

We notice from LeNet-STD results in Fig. 9.2 that sensitivity increases from 50% to 80% with oversampling, undersampling, lead shift-oversampling, time reversing-undersampling and time reversing-oversampling while maintaining a high AUC (80%). The accuracy and specificity values are maintained at 90% except for undersampling and time reversing-undersampling where it decreases to 70%. Time shift does not enhance sensitivity even when combined with oversampling and undersampling. The overall performance of the CNN is significantly better than that of MLR, especially in terms of sensitivity and AUC. The classification performance with time shift is poor with a sensitivity not exceeding 55% with both LeNet-STD and MLR. This DA method seems to hide STD information needed for an accurate training.

The bootstrapping statistical tool is used to quantify the uncertainty of each trained

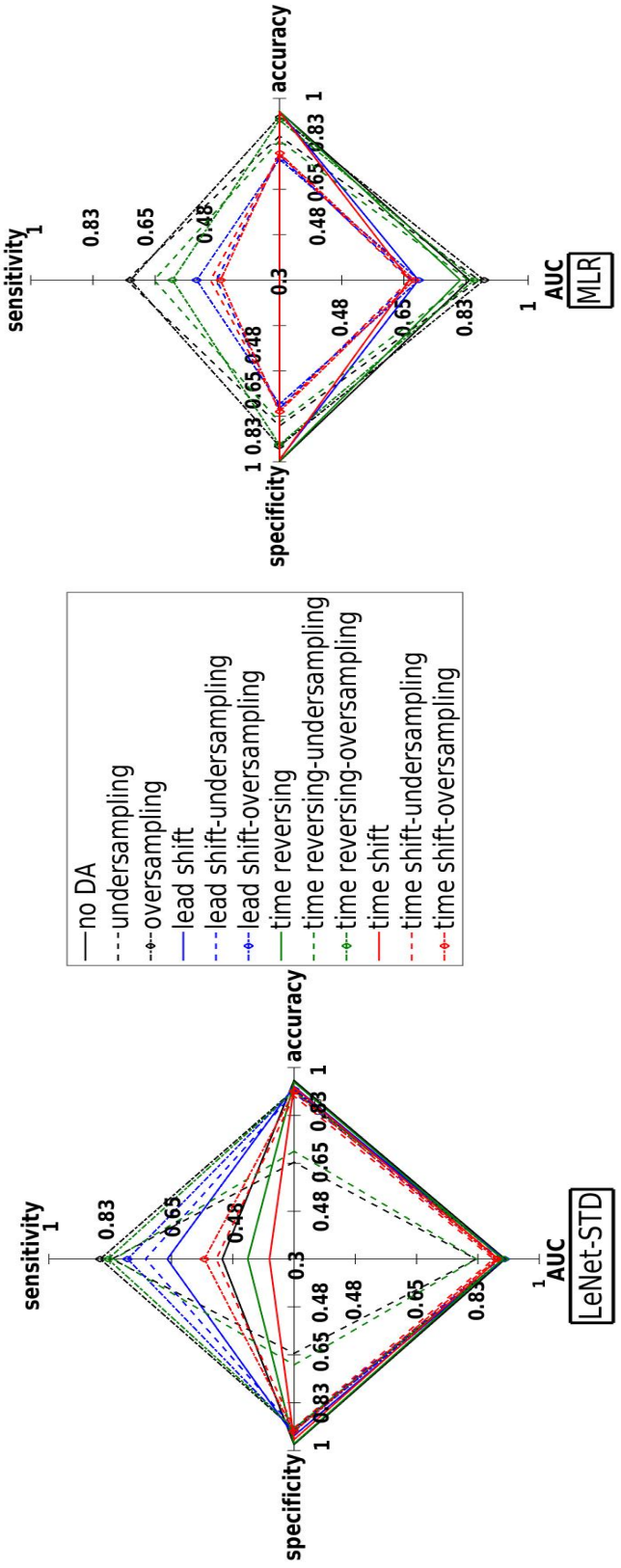

Figure 6.5: Average classification performance on test set with data augmentation over bootstrapping.

classifier. First, 100 subdatasets of 300 samples (about 10% of the test set) are selected at random with replacement. Second, the model performance is calculated for each subdataset then the metrics are averaged. We notice that the means coincide perfectly with the metrics of the original test set and that the standard deviation is zero, showing the robustness of the trained classifiers.

DA	**CIR** %
no DA	5
undersampling or oversampling	100
lead shift	51
time reversing	10
time shift	56

Table 6.2: Class imbalance ratio (CIR) calculated for the training datasets.

6.5 Summary and Conclusions

In our problem, DA aims as a key ingredient to improve the identification accuracy of STD bipoles, thus recognizing atrial areas potentially driving persistent AF to be targeted in ablation with a high degree of accuracy. High dataset imbalance and insufficient amount of STD samples leads to poor classification results, motivating the need for DA techniques. Among the DA techniques considered, oversampling performs well whether considered alone or combined with lead shift or time reversing. With LeNet-STD, sensitivity increases from 50% to 80% while maintaining high AUC and accuracy values at 90% and slightly decreasing specificity. The overall performance of LeNet-STD is better than that of MLR which can be explained by the expressive power of CNNs in addition to the fact that the training dataset is large and needs a large model in order to be correctly classified. Bootstrapping showed that the trained classifiers are robust with respect to the test dataset. The benchmarking of DA methods can be extrapolated to further biomedical signal classification problems like ECGs and electroencephalograms.

The next chapter will analyze statistical and handcrafted features in addition to the benchmarking of several data representations and classifications algorithms in order to automatically identify STD EGMs.

Chapter 7

Automatic Identification of Spatiotemporal Dispersion Using the VAVp Time Series

Contents

7.1 Introduction

The previous chapter proposed to handle the imbalance between the number of STD to the non-STD samples with the use of adapted DA techniques in order to enhance the classification performance of baseline classifiers such as MLR. Oue preliminary classification models require a labeled dataset of raw EGM recordings and are able to learn

the decision rule automatically through the training process with mini-batch GD. The present chapter aims to automatically characterize and identify ablation sites in STD-based ablation differently.

Numerical simulations in the reviewed paper [9] suggest that the distribution of the time series of maximal voltage absolute values at any of the PentaRay bipoles (VAVp) is significantly affected by the dispersion pattern. Motivated by this hypothesis, we suggest to consider VAVp as a key feature for STD identification. Two different approaches are investigated. First, engineered and statistical features are extracted manually from VAVp time series and fed to supervised classifiers. This first approach is called *feature engineering*. In particular, we study the kurtosis [95] and distribution (histogram) of VAVp time series. Synthetic multichannel EGMs are generated to mimic the STD pattern and to gradually evaluate its effect on VAVp statistics. The histograms of VAVp samples from both STD and non-STD real datasets are also clustered using HCA [48]. Experiments are conducted on synthetic signals, through several Monte Carlo (MC) simulations, and then on real datasets.

We find in the literature several studies that characterize the AF pattern with statistical measures and handcrafted features such as [110]. The study characterizes the nature of the localized atrial activation during AF in 34 patients, following open heart surgery based on the analysis of bipolar atrial electrograms. The features used include the signal morphologies, shapes, polarities, amplitudes and beat-to-beat intervals. Results show that these characteristics differ from one AF type to other. As mentioned in Sec. 2.5.4, AF can be classified in several stages according to the duration and frequency of abnormal episodes. The study [111] uses multichannel atrial EGM recordings (244 unipolar electrode) for the classification of atrial fibrillation into three stages. The classification is based on a measure of the degree of complexity of atrial activation during AF. The methods proposed in [110] and [111] are designed to characterize global AF patterns whereas the present doctoral thesis focuses on STD which is a local AF pattern.

The second approach is called *automatic feature extraction*. It consists in classifying raw VAVp recordings using supervised ML tools. Both approaches, feature engineering and automatic feature extraction, are complementary. On the one hand, the statistical analysis of engineered features is intuitive and easily interpretable although results are strongly dependent on an appropriate choice of features. On the other hand, even successful modern ML tools like CNN can automatically extract classification features from raw data through its deepest layers. However, these features are not easily interpretable and validated from a medical perspective.

7.2 VAVp Times Series

The VAVp signal is a compact (1D) representation of multichannel recordings introduced in [9]. Let $\mathbb{X}$ represent an EGM sample and $\mathbb{X}(i,j)$ the j^{th} time sample of the i^{th} PentaRay bipole. As shown in Fig. 7.1, the VAVp time series is calculated as follows:

1. First, the VAV matrix contains in its rows the absolute values of each channel of the multilead EGM recording:

$$\text{VAV}(i,j) = |\mathbb{X}(i,j)|$$

2. Second, the 1D VAVp signal is computed as the maximal values at each time sample of VAV matrix over the leads (rows) dimension:

$$\text{VAVp}(j) = \max_i \text{VAV}(i,j)$$

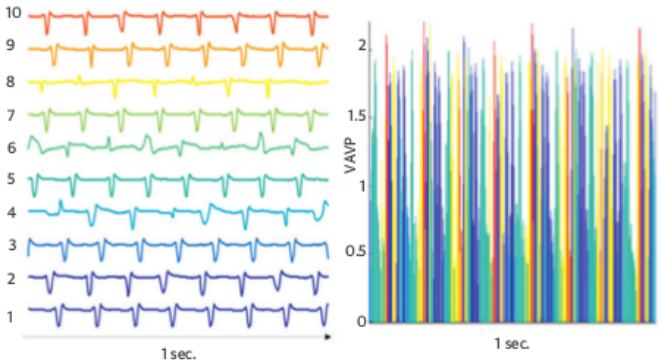

Figure 7.1: Plot of multichannel EGM recordings (Left) and the corresponding VAVp time series (Right). Image modified from [9].

Numerical simulations in [9] show that the histograms of VAVp distribution, that we refer to as $h(\text{VAVp})$, depend on STD: $h(\text{VAVp})$ is peaky and concentrated around zero if the virtual PentaRay is positioned in non-STD areas, at the periphery of a driver with slow excitation for instance, but it gets more spread for EGMs recorded in STD areas of the atria, like rotors as shown in Fig. 7.2. Thus, it is hypothesized that VAVp could be used for STD identification. The present chapter aims at testing this hypothesis.

7.3 Methods

This section recalls briefly statistical measures, classification algorithms and performance metrics used in the experiments conducted in the present chapter and whose

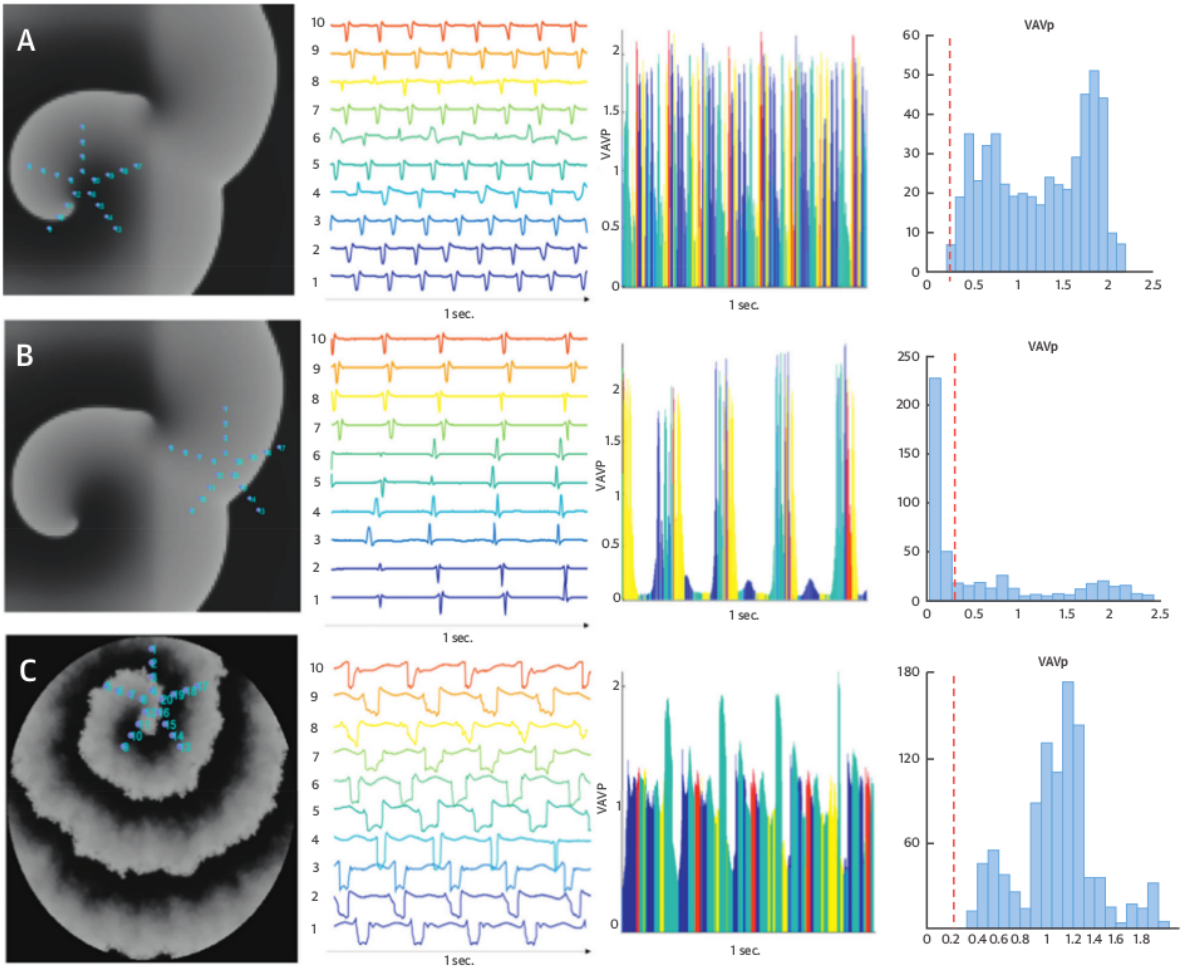

Figure 7.2: VAVp distribution (**A**) at the center of the driver, reminiscent of patients' dispersion areas, (**B**) at the periphery of the driver and (**C**) in the interstitial fibrosis condition. This result is obtained with the aid of numerical simulations. Image taken from [9]

detailed definitions are provided in Chap. 4.

7.3.1 Synthetic VAVp Time Series

We recall that a realistic data model for generating synthetic multichannel EGM recordings containing parametrized STD patterns is presented in Sec. 5.3. Each lead is modeled as a concatenation of single sawtooth patterns that have different amplitudes. The amplitude is linked to the fractionation degree d. The model is parametrized with different hyperparameters that characterize the STD pattern, such as the number of consecutive delayed leads N_{delay} and the number of fractionated leads N_{frac}. Once a synthetic EGM is generated, its VAVp distribution is calculated similarly to an original EGM record.

7.3.2 Skewness and Kurtosis

Skewness is a measure of the asymmetry of a variable's probability distribution around its mean. We compute unbiased estimates of kurtosis and skewness [95]. In statistics, kurtosis refers to a measure of peakedness of a distribution. It quantifies whether the data are heavy-tailed (high κ_4) or light-tailed (low κ_4) relative to a normal distribution ($\kappa_4 = 0$). Further details about skewness and kurtosis are given in Sec. 4.6.

7.3.3 Histogram Clustering

Given a set of data samples, HCA [48] partitions them into agglomerations using a homogeneity criterion, so that points within each cluster are similar and points from different clusters are dissimilar. As explained in Sec. 4.3.2, HCA treats each observation as a singleton group at the beginning. Then it merges pairs of groups until all clusters are agglomerated. The linkage can be presented by a dendrogram. A cutoff is then chosen and the clusters are formed [48]. The most important advantage of using HCA is that it does not require a previous specification of the number of clusters which is in general unknown, unlike the k-means algorithm [47], as explained in Sec. 4.3.1.

7.3.4 Classification Algorithms and Metrics

Supervised ML techniques, namely LDA [53], SVM [57] and a shallow CNN [70], are trained on labeled datasets to classify samples from STD and non-STD categories. These classifier used as baseline models for 1D signal analysis. The mathematical formulations of classification algorithms are given in sections Sec. 4.4.2, Sec. 4.4.4 and Sec. 4.4.6, respectively. Each classifier learns a classification rule then applies it to new data. The performance of ML models is evaluated through the following metrics: 1) accuracy; 2)

area under the receiver operating characteristic curve (AUC); 3) sensitivity (true positive rate), measuring the proportion of actual STD samples that are correctly identified as such; 4) specificity (true negative rate), measuring the proportion of actual non-STD samples that are correctly identified as such. More details about the performance metrics are provided in Sec. 4.5.

7.4 Experiments

To test the validity of the VAVp time series for STD identification, a set of experiments is carried out. Experiments including the analysis of VAVp statistics, histograms clustering and classification belong to the feature engineering approach, whereas the automatic feature extraction approach consists in classifying the raw VAVp signals directly.

7.4.1 VAVp of Synthetic Data

As a first attempt, we investigate the effect of the number of consecutive delayed channels (N_{delay}) on the VAVp distribution. Fig. 7.3 shows that as N_{delay} grows, the kurtosis decreases which means that the distribution becomes flatter. In absence of fractionation ($d = 1$) and for $N_{\text{delay}} = \{1, 2, 3, 4\}$ we have $\kappa_4 > 0$, the VAVp histogram approximates a normal distribution ($\kappa_4 \approx 0$) for $N_{\text{delay}} = \{5, 6\}$ but the kurtosis becomes negative for $N_{\text{delay}} = \{7, 8, 9\}$. This experiment is repeated for different delay values in the set $\{40, 70, 100\}$ ms. The resulting curves almost coincide, which means that delay range is not significant.

The combined effect of fractionation and delay of contiguous leads is assessed next. For each N_{delay}, 100 MC simulations are conducted. For each MC run, both the number (N_{frac}) of fractionated leads and their positions are selected at random. Then, VAVp distribution statistics are computed. Finally, the average values of kurtosis and skewness are obtained. The curves in Fig. 7.3 show the resulting kurtosis for $d = 8$ and $d = 16$ over the 100 MC runs. The values are presented as mean $\pm$ standard deviation. We notice that similar behavior is obtained for $d = 1$ and $d = 8$. However for more fractionated signals ($d = 16$) the kurtosis decreases from 4 to 0. The distribution passes smoothly from peaky to Gaussian as N_{delay} grows. Experiments also demonstrate that skewness is a decreasing function of d which means that the VAVp distribution gets symmetric around its mean, as d grows. Besides, skewness is much less sensitive to N_{delay} compared to kurtosis. For this reason we decide not to consider it in our study.

In order to asses the effect of fractionation on VAVp distribution, fractionated multichannel EGMs are generated. Each synthetic matrix contains a growing number of fractionated leads $N_{\text{frac}} \in [1, 10]$. Fractionation degrees (d) are divided into three ranges

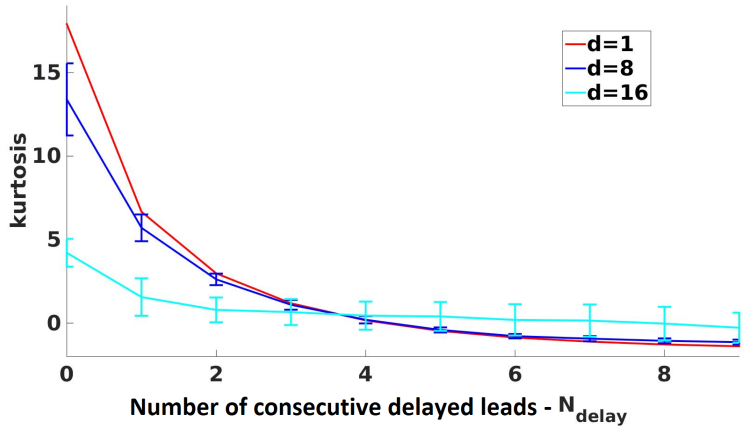

Figure 7.3: Effect of the delay on VAVp distribution. The fraction degree is d and the number of consecutive delayed leads is N_{delay}.

$\{[1,6],[1,12],[1,19]\}$. For each range and for each N_{frac}, a series of 100 MC simulations is conducted where the fractionation degrees are chosen randomly in that range. The values of kurtosis are computed as in the previously described experiment. Fig. 7.4 shows that the kurtosis is a decreasing function of both N_{frac} and d. The effect of N_{frac} on the distribution is less important than d regarding the slope of the curves (Fig. 7.4).

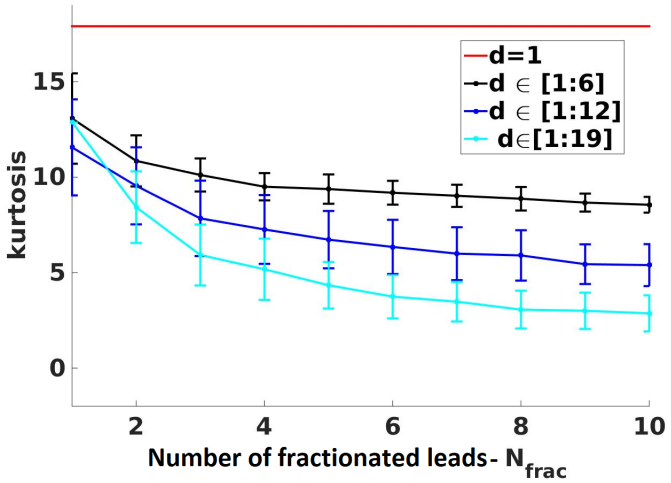

Figure 7.4: Combined effect of fractionation on VAVp distribution. The fraction degree is d and the number of fractionated leads is N_{frac}.

To summarize, VAVp distribution gets flatter and its tail lighter as the dispersion patterns represented by the number of delayed leads, fractionation degree and fractionated lead positions are emphasized whether simultaneously or separately. The results on synthetic signals confirm the hypothesis that the VAVp distribution is significantly affected by the STD pattern and motivates the idea of automatically characterizing STD

EGMs with the use of their VAVp distributions.

7.4.2 VAVp of Real Data

The real dataset exported from Nice CHU includes a cohort of 23084 points composed of 1140 STD and 21944 non-STD samples, belonging to the LA cartographies. The ratio of STD to non-STD samples (CIR) is equal to 5%. We recall that for each cartography of each patient, 10 EGM bipoles are recorded per location and each acquisition can be stored in a data matrix of dimensions 10×2500.

Kurtosis Analysis

Investigating the effect of STD pattern using real EGM data is more challenging than in synthetic signals generated in controlled conditions. We compute the kurtosis of multichannel EGM samples belonging to both STD and non-STD datasets. The computed values lie in the following intervals: $\kappa_{4_{\text{STD}}} \in [2, 410]$ while $\kappa_{4_{\text{non-STD}}} \in [1, 621]$. Mean and standard deviations values are $\kappa_{4_{\text{STD}}} = 23 \pm 34$ and $\kappa_{4_{\text{non-STD}}} = 29 \pm 41$, which means that in average both VAVp distributions have sharp histograms but the STD is slightly flatter than the non-STD one, as illustrated in the boxplot in Fig. 7.5. As a result, we conclude that looking at metrics related to the VAVp histograms like kurtosis is not a significant hand-engineered feature in STD identification.

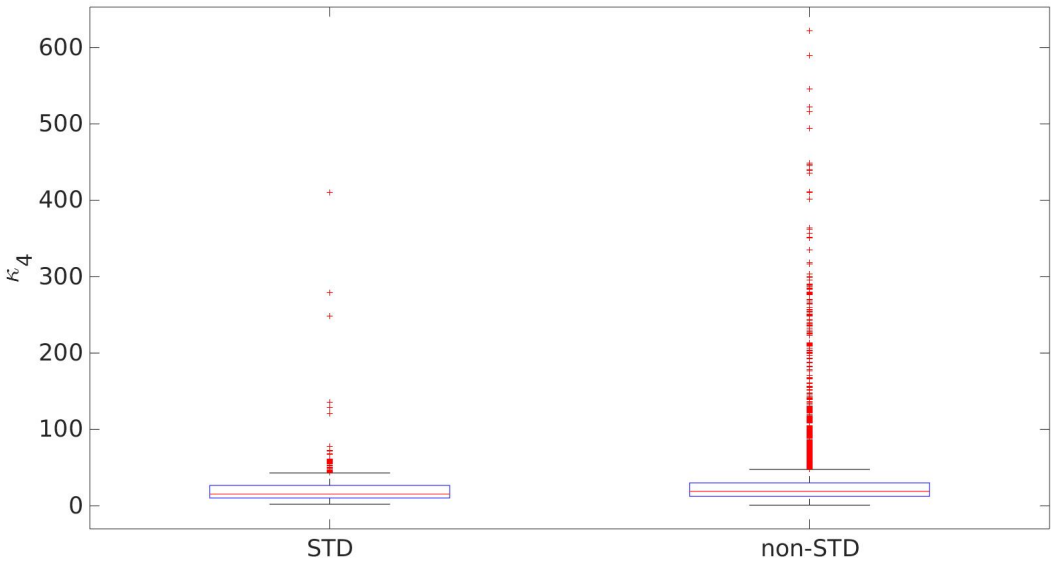

Figure 7.5: Boxplot of the kurtosis of STD and non-STD samples.

Histogram Clustering and Classification

We attempt to analyze the VAVp histograms coming from both STD ($h(\text{VAVp})_{\text{STD}}$) and non-STD ($h(\text{VAVp})_{\text{non-STD}}$) datasets. We first cluster the histograms of each class separately using the HCA [48] algorithm. We opt for Hellinger distance, introduced in Sec.4.3.2, as a homogeneity criterion because it is a recommended metric to measure the similarity between sparse data like sparse histograms [49]. Cutoff values are chosen visually from the dendrograms. We opt for 6 STD and 10 non-STD clusters of $h(\text{VAVp})$ then we compute the averaged histograms of the representative element of each cluster. No obvious dissimilarity in terms of histogram sharpness nor flatness is obtained by comparing STD to non-STD clusters as shown in Fig. 7.6. For instance, clusters 4 and 5 belonging respectively to STD and non-STD categories are both sharp, reflecting a synchronized cardiac activation between the channels of the EGM sample. Similarly, clusters 3 and 8 belonging respectively to STD and non-STD categories are both flat, reflecting a non-synchronized cardiac activation between the EGM channels. This finding rejects the hypothesis that VAVp distribution is a key variable in STD identification, as confirmed by both experiments on simulated data in Sec. 7.4.1 added to Figs. 7 and 8 in the medical reference [9].

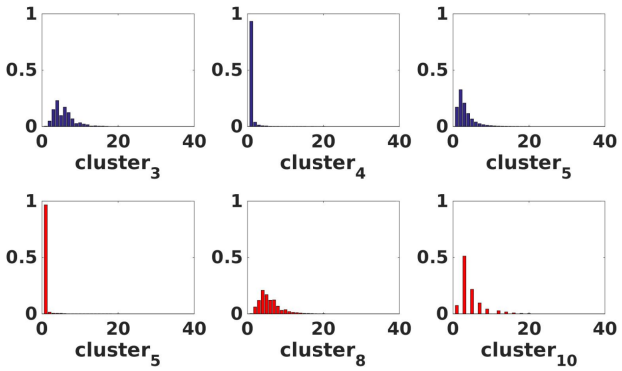

Figure 7.6: VAVp histograms of the STD (**blue**) and non-STD (**red**) clusters.

To further confirm this finding, we try to classify the labeled histograms into STD vs. non-STD using ML tools like LDA, SVM with both linear and Gaussian filter and a shallow CNN but the classification performance is very poor. Similar results are obtained with both balanced and non balanced training sets. This result may be explained by the two following reasons:

- First, the non-STD class contains a wide range of multichannel EGM types that certainly have different patterns that would result in a wide range of kurtosis and VAVp distributions.
- Second, the labels of STD samples do not contain detailed information about the

number of fragmented (CFAE) leads neither the positions of delayed leads. As a result, the STD dataset can be assimilated to a bag containing samples with all possibilities of delay values, N_{delay}, N_{frac} and d.

Raw VAVp Classification

Motivated by the poor results obtained with the analysis of VAVp distribution, we propose in this section to analyze the raw VAVp time series. To this end, we train supervised ML models with the labeled dataset. The classifiers used are:

- LDA
- SVM with a Gaussian kernel, as preliminary experiments showed that the performance of a Gaussian kernel is superior to the linear one
- Shallow CNN composed of a 1D conv layer with 32 nodes followed by a dropout layer with a dropout probability equal to 0.2 then a FC layer with a sigmoid (linear regression) activation function. As explained in Sec. 4.4.5, dropout is a regularization technique used to avoid overfitting [68]. We tried to add several types of pooling layers (max, average and global average) to the CNN architecture but they all degraded the classifiers's performance.

Regarding the low CIR (5%) and lack of a sufficient amount of STD samples, we propose to apply DA to the training set. DA consists in synthesizing new samples belonging to the minority class by applying transformations to its original samples. Based on our comparative study detailed in Chap. 6, ROS remains a simple but effective DA option for STD identification in the same dataset. ROS consists in forming a balanced superdataset by replicating randomly samples from the minority class (STD) until they reach the number of the majority class (non-STD), as defined in Sec. 6.3.2.

In order to handle overfitting and asses the ability of the trained ML models to generalize the classification rule, we use 5-fold CV technique [92]. A rule of thumb is to choose k equal to 5 or 10 [93]. Regarding the size of our dataset, we opt for 5-fold CV. Hence each fold (formed randomly) contains a sufficient number of original STD samples, besides the synthetic ones obtained with ROS. In order to train the CNN, we use ADAM optimizer [66] with validation AUC as an early stopping criterion. In each CV round, the test dataset is partitioned into two equal-sized subsets that will form the new validation and test sets. This guarantees that the model does not see the test samples during the training phase. Classification results on the test set are given in Tab. 7.1. The values of accuracy, AUC, sensitivity and specificity are around 90% and present a very low variability of the order of 10^{-3}. However, we notice that the performance of LDA is inferior to that of non linear classifiers (SVM and CNN). For instance, the AUC

value of the CNN is 96% while it is equal to 90% for the SVM and 87% for the LDA. The AUC informs about the combined true classification rates of both STD and non-STD classes. The superiority of the CNN may be explained by the ability of this model to automatically extract key classification features through its conv filters.

	Accuracy	**AUC**	**Sensitivity**	**Specificity**
LDA	0.883	0.866	0.848	0.885
SVM	0.927	0.905	0.880	0.930
CNN	0.917	0.964	0.867	0.922

Table 7.1: Average VAVp time series classification performance on test set through 5-fold CV.

7.5 Summary and Conclusions

Motivated by the hypothesis in [9] that the distribution of VAVp signals is affected by the STD pattern, we propose to automatically identify STD areas in multichannel EGM recordings in persistent AF through two complementary approaches based on the VAVp time series. Preliminary experiments using synthetic signals in controlled conditions demonstrate that VAVp statistical features like kurtosis are indeed good indicators of STD in AF multipolar EGMs. However, further experiments in real AF data show that these features do not allow STD identification. Likewise, VAVp histograms of STD and non-STD datasets clustered using HCA do not present significant dissimilarities. Modern ML tools that automatically identify characteristic features from the raw VAVp time series prove more efficient, with AUC up to 96% for CNN.

In the next chapter, we will benchmark different classification models composed of several data formats, including VAVp, and suitable algorithms for the classification of EGMs into STD vs .non-STD. The size of the database will also be increased in order to provide more relevant clinical results.

Chapter 8

Comparative Study of Several STD EGM Classification Models

Contents

8.1 Introduction

As introduced in Sec. 1.2 and highlighted in Chaps. 6 and 7, the main objective of the preset doctoral thesis is to overcome the limitations of the visual inspection of STD EGMs by designing a decision-aid solution that helps interventional cardiologists detect STD patterns automatically.

Our preliminary results in a previous study on automatic detection of STD from multichannel EGM recordings were promising, as demonstrated in Chap. 6. The classification performance on the test dataset reached 90% of accuracy and 80% of AUC using shallow CNN, called LeNet-STD. But values of precision are low because of the

lack of STD samples. We also addressed the issue of class imbalance and lack of training samples from STD class through adapted DA methods and concluded that ROS method yields the best performance. Chap. 7 presented a complementary recent study where we addressed the classification of handcrafted features from EGM data. As suggested in [9], we studied a time series computed from multichannel EGM signals to perform STD detection. The time series is called VAVp. Experiments showed that classifying EGM samples in the VAVp format with baseline algorithms such as a shallow CNN or LDA yield a good performance when combined with ROS. However, the characterization of the STD pattern from the VAVp time series on the basis of handcrafted features and statistical measures such as kurtosis and distributions analysis were unsuccessful, when dealing with real EGM recordings.

Based on these preliminary findings, the present study seeks the most optimal EGM classification model, composed of a data representation and an adapted classification algorithm, in order to detect STD locations with the highest performance and lowest computational cost in terms of prediction time and resources. The study dataset includes a cohort of over 35000 10-channel EGM signals acquired from different cartographies of persistent AF patients. The dataset is described in detail in Sec. 5.2.

This contribution completes and consolidates our preliminary results in Chaps. 6 and 7 by putting forward a feature selection technique for a better optimized EGM classification solution, including a bigger dataset (35000 compared to 23000 samples). For this purpose, we design and benchmark different data representations:

1. Raw EGMs stored in matrices
2. 2D images obtained by subplotting the curves of the 10 leads one under the other, as it is currently done by the mapping system in the hospital
3. 3D image, tensor as a result of stacking three shifted images of EGM curves
4. 1D VAVp time series used as a compact representation of the multichannel recordings, presented in Chap. 7

The data representations are combined with suitable classifiers including MLR [56], dimensionality reduction with PCA [50] followed by SVM [58], CNN [79], TL [10] of VGG16 model [81]. The originality of this contribution is two-fold:

1. Different data representations are benchmarked and TL is applied for the first time to the identification of dispersion pattern in multichannel EGMs
2. Our decision-aid solution can be implemented in real time with moderate computational resources, thus potentially improving catheter ablation success rates while

reducing the duration and costs of the STD-based ablation interventions of persistent AF

8.2 Data Representations

Starting from the basic matrix format presented earlier in this thesis, data can be reformatted in different ways to match different classification algorithms:

1. **Matrix**: The first classification scheme uses basic matrices of dimensions 12×2500, the result of applying the circularity transformation on raw recordings of the 10-lead EGMs, as explained in Sec. 5.2.4. Matrices can be treated as images with CNN models, as proposed in [112] for a human activity recognition task.

2. **2D image**: The image format consists in subplotting the curves of the 10 leads, one under the other, exactly as it is currently done by the CARTO system during the mapping phase. This 2D plot of EGMs is used in the hospital and represents the most easily understandable representation by the eye of an interventional cardiologist. We will check in the present study if this format is also convenient for ML classifiers.

3. **3D image**: This format enables the use of CNN models conceived to train on 3D RGB images like in ImageNet [82] dataset (natural colored images). Hence, this data format enables the use of TL [10]. The 3D tensor is the result of stacking three 2D images along the depth dimension. The first slice (image) is a subplot of the ordered leads $\{\ell_1, \ell_2, \ell_3, \ell_4, \ell_5, \ell_6, \ell_7, \ell_8, \ell_9, \ell_{10}\}$, similarly to 2D image format. The second slice is the subplot of the leads shifted in a circular way. The leads order becomes $\{\ell_2, \ell_3, \ell_4, \ell_5, \ell_6, \ell_7, \ell_8, \ell_9, \ell_{10}, \ell_1\}$. The third slice represents a subplots of leads shifted twice as follows $\{\ell_3, \ell_4, \ell_5, \ell_6, \ell_7, \ell_8, \ell_9, \ell_{10}, \ell_1, \ell_2\}$. Similarly to the circularity transformation, the idea comes from the rotational symmetry of the PentaRay branches. The interest in shifting the leads remains in allowing the 3D filter to systematically analyze neighboring leads along the depth dimension as it already does along the spatial dimensions (matrix slice). For instance the first convolutional window captures leads $\{\ell_1, \ell_2, \ell_3\}$ across the three slices. This data format allows us to represent EGMs in a 3D tensor with a depth of three thus mimicking the shape of RGB images and enabling the use of algorithms designed for this input shape.

4. **1D signal**: VAVp times series is a compact representation of multichannel recordings that has shown promising results in Chap. 7. The calculation of VAVp signal is

provided in Sec. 7.2. Preliminary experiments showed that the supervised classification of raw VAVp time series itself into STD vs. non-STD categories is promising. Results on the test are good with values of accuracy, AUC, sensitivity and specificity around 90% with low variability (10^{-3}). Complementary metrics such as precision and F1 score are also computed and analyzed in the present work. In addition, a larger dataset of 35000 EGMs are included in the present chapter compared to 23000 in Chap. 7.

8.3 EGM Classification Models

This section reviews some state-of-the-art studies about the classification of cardiac signals using ML and statistical analysis, then introduces the models we propose for the automatic classification of multichannel EGM records in STD-based ablation. Finally, we explain our methodology for training ConvAC algorithms as a tensor decomposition and present preliminary results obtained on synthetic data (toy example).

8.3.1 State-of-the-Art Models

The topic of classifying multichannel EGM recordings in AF patients for the automatic identification of STD sites is relatively recent. To our knowledge, no public studies dealing with the same topic were available at the time of writing the present thesis. However, a wide variety of studies about the automatic classification of ECG and EGM signals in AF using ML and DL can be found in the literature. Deep CNN architectures are adapted to the classification of 12-lead ECG recordings. For instance, studies [115] and [P5] propose a CNN inspired by the VGG architecture, defined in Sec. 4.4.6, to perform the multilabel classification of PhysioNet Challenge datasets (2018 and 2020, respectively) [108] into different arrhythmias including AF in addition to NSR. The architecture designed in [115] implements LSTM and attention module [116] added to conv ones, as illustrated in Fig. 8.1. Hence, the model yields an average F1 score of 81.2% over all rhythm classes. The study in [117] proposes an automatic and fast arrhythmia classifier based on a simple brain-inspired ML approach from the analysis of a single ECG lead. The solution consists in two steps:

1. Processing of the ECG recordings through filtering, detecting heartbeats, segmenting heartbeat and extracting features.

2. Classification of the preprocessed ECG recordings into two classes using an ensemble of echo state networks.

The classification criterion is based on the morphology of heartbeats. Experiments on

the MIT-BIH Arrhythmia [118] database provide high values of sensitivity and positive predictive rate, respectively equal to 92.7% and of 86.1%. The MIT-BIH Arrhythmia database consists of 48 half-hour ECG recordings sampled at 360 Hz. Each ECG record has two channels. Samples are annotated into 5 classes, however the experiments in [117] perform a binary classification by fusing the different classes into 2 categories only.

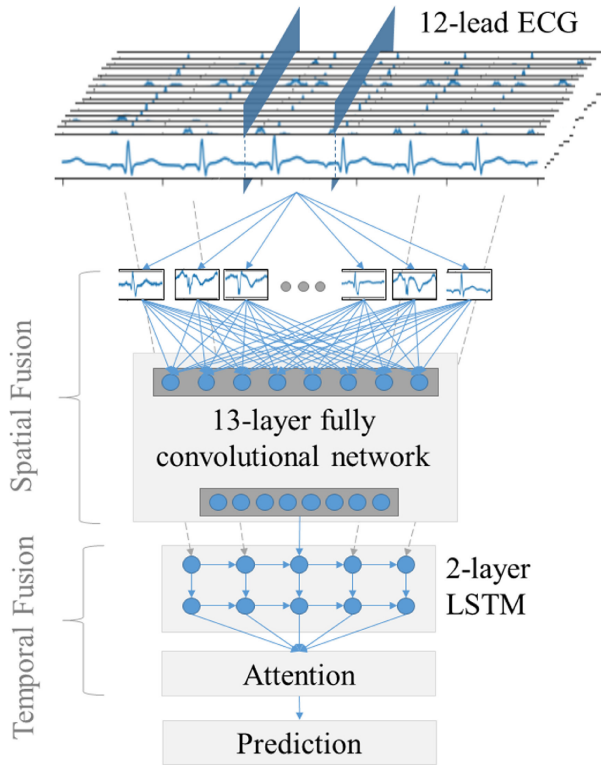

Figure 8.1: DNN model proposed for the classification of multichannel ECG recordings acquired from PhysioNet challenge 2018. The 12-lead ECG is fed into fully convolutional layers and therefore generate a preprocessed time series. This time series is further fed into LSTM cells to exchange information among different time points. An attention module accepts output of LSTM cells, assigns weights for different time points and outputs a final result. Image taken from [115].

Classifying 64-channel EGM recording into AF vs. atrial flutter, based on their shape rather than rate and regularity is studied in [119]. The study implements a deep CNN and yields an average accuracy of 80% over 10-fold CV for a population of 86 patients (25 female, age 65±11 years) in persistent AF. Each EGM recordings spans 60 s.

The architectures proposed in [117] and [119] area can be adopted for the classification of our multichannel EGM dataset in future works, as they require a huge computational resources and training time. A trade-off choice in the present thesis consists in using the VGG16 CNN architecture for TL.

Contrarily to the previously described end-to-end learning approaches, other studies propose to extract features from EGM data then use these characteristics as the input of classifiers. For instance, the goal of the study [120] is to identify arrhythmogenic sources maintaining AF based on the analysis of features extracted from real EGM records. A genetic algorithm is used to select the optimal subset of features out the 92 calculated originally, the k-nearest neighbors algorithm is used to classify data into four groups (according to the fractionation levels). The classification performance obtained on a dataset of 429 real EGM records achieves a sensitivity of 90% and specificity of 97%.

8.3.2 Proposed Models

The association between the different data formats and classifications algorithms studied in the present chapter is schematized in Fig. 8.2. A data model corresponds to the coupling between a data representation presented in Sec. 8.2 and a suitable classifier and can be summarized as follows:

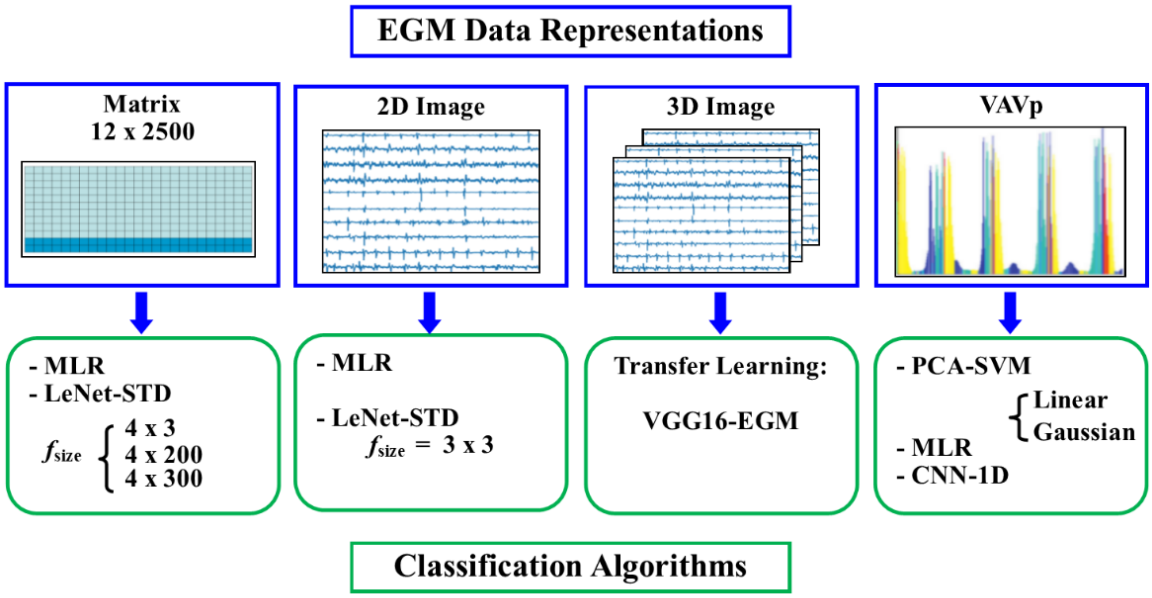

Figure 8.2: Classification models adapted to the different EGM data formats considered in this study. A classification model is obtained as the combination of of a data representation and an adapted classification algorithm.

Matrix Classification

MLR, LeNet-STD and a shallow ConvAC are used for the classification of EGM matrices. LeNet-STD was introduced in Chap. 6, Sec. 6.2 as a proof of concept that 2D CNN

architectures are adapted to STD detection. The interest in shallow ConvAC networks lies in their equivalence to a tensor decomposition. We will try to take profit from this equivalence to train a ConvAC network by solving its algebraic equation. The procedure will be detailed in Sec. 8.3.3. For LeNet-STD, we study the effect of the receptive field (filter size) of the network's first conv layer, denoted (f_{size}). We recall that the STD pattern is usually observed on a minimum of three bipoles positioned on two adjacent PentaRay branches and the average AFCL value is typically 200 ms. A natural choice is $f_{\text{size}} = 4 \times \alpha_1 \text{AFCL}$ ($\alpha_1 \geq 1 \in \mathbb{Q}$) with $\alpha_1 = \frac{3}{2}$ as proposed in 6, thus capturing cardiac activation along a minimum of one and a half AFCL. This yields $f_{\text{size}} = 4 \times 300$. The model is denoted LeNet-STD$_{(4\times300)}$. We investigate further settings such as $f_{\text{size}} = 4 \times 3$ and $f_{\text{size}} = 4 \times 200$. The models are denoted LeNet-STD$_{(4\times3)}$ and LeNet-STD$_{(4\times200)}$, respectively. A classical window size is 3×3 [81] in image classification tasks, whereas f_{size} equal to 4×200 is conceived to capture exactly an entire AFCL.

2D Image Classification

MLR and LeNet-STD$_{(3\times3)}$ are used for images (EGM plot) classification, with f_{size} chosen to be 3×3 as a typical receptive field for image classification. In deed such a filter size is the smallest receptive field to capture the patterns of left/right, up/down and center, as explained in [81].

3D Image Classification

We perform the classification of 3D images with a deep CNN called VGG16-EGM. VGG16-EGM is the result of transferring the VGG16 architecture. Indeed, we re-use the first trained conv blocks of the VGG16 model. The transferred layers serve as a feature extractor. Their output is fed to three consecutive FC layers with 1024, 128 and 2 nodes, respectively. Only FC layers have learnable weights. Hence, VGG16-EGM is adapted to 3D image classification and can be trained on the EGM dataset.

VAVp Classification

In addition to MLR, SVM is another baseline linear classifier in ML. However, classifying VAVp time series is not possible with SVM because this input has 2500 covariates and SVM is not adapted to a such huge amount of variables. We propose then to reduce the dimensions of VAVp signals with PCA, as explained in Sec. 4.4.1. We use the hard threshold technique at 95% (of explained variance) to decide the number of the first PCs to keep (r). Then, we feed the reduced version of VAVp to SVM. Linear and Gaussian kernels are benchmarked. The models are denoted respectively PCA-SVM$_{\text{Lin}}$ and PCA-SVM$_{\text{Gauss}}$. We also perform classification with a shallow 1D CNN (CNN-1D).

8.3.3 Training ConvAC as a Tensor Decomposition

As a first step, we suggest to classify EGM signals using a shallow ConvAC [84], theoretically proven to be expressively efficient as mentioned in Sec. 4.4.6. In order to train the ConvAC, while avoiding the pitfalls of gradient descent approach, we propose to model the mapping function of the network as a high-dimensional tensor and estimate the optimal network parameters by solving the algebraic solution of its CPD decomposition [103] (Sec. 4.9).

The ConvAC architecture is described in Fig. 4.20. Let $\mathbf{X}$ be an instance image divided into N patches $\{(\tilde{x}_i)\}_{1,\dots N}$, a set of M representations functions $f_{\theta_1} \dots f_{\theta_M} : \mathbb{R}^s \to \mathbb{R}$ are applied to the patches. The set $\{f_{\theta_d}\}_{d\in\{1,\dots,M\}}$ should be total and linearly independent in $L^2(\mathbb{R}^s)$. The transformed image $\{f_{\theta_d}(\tilde{x}_i)\},\ i \in \{1,\dots,N\},\ d = \{1,\dots,M\}$ is then passed to a convolution layer with linear filters of size 1×1, performing the scalar product with the representation function outputs:

$$\sum_{d=1}^{M} a_d^{z,i} f_{\theta_d}(\tilde{x}_i),$$

after which a product pooling is applied to the Z feature maps:

$$\prod_{i=1}^{N} \Big(\sum_{d=1}^{M} a_d^{z,i} f_{\theta_d}(\tilde{x}_i) \Big).$$

A final dense layer outputs the score function $h_y, y \in \{1,\dots,Y\}$:

$$h_y(\tilde{x}_1,\dots,\tilde{x}_N) = \sum_{z=1}^{Z} a_z^y \prod_{i=1}^{N} \Big(\sum_{d=1}^{M} a_d^{z,i} f_{\theta_d}(\tilde{x}_i) \Big).$$

Class y^* is attributed to X via maximization of per-label h_y.

$$\begin{aligned}
h_y(\tilde{x}_1,\dots,\tilde{x}_N) &= \sum_{z=1}^{Z} a_z^y \prod_{i=1}^{N} \Big(\sum_{d=1}^{M} a_d^{z,i} f_{\theta_d}(\tilde{x}_i) \Big) && (8.1)\\
&= \sum_{z=1}^{Z} a_z^y \Big(\sum_{d_1=1}^{M} \mathbf{a}_{d_1}^{z,1} f_{\theta_{d_1}}(x_1) \Big)_1 \dots \Big(\sum_{d_N=1}^{M} \mathbf{a}_{d_N}^{z,N} f_{\theta_{d_N}}(x_N) \Big)_N && (8.2)\\
&= \sum_{d_1,d_2\dots,d_N}^{M} \mathcal{A}^y_{d_1,\dots,d_N} \prod_{i=1}^{N} f_{\theta_{d_i}}(\tilde{x}_i) && (8.3)\\
&= < \mathcal{A}^y, \mathcal{F} > && (8.4)
\end{aligned}$$

where

$$\mathcal{A}^{y}_{d_1,\ldots,d_N} = \sum_{z=1}^{Z} a^{y}_{z} \mathbf{a}^{z,1}_{d_1} \ldots \mathbf{a}^{z,N}_{d_N} \tag{8.5}$$

and

$$\mathcal{F}_{d_1,\ldots,d_N}(\tilde{x}_1,\ldots,\tilde{x}_N) = \prod_{i=1}^{N} f_{\theta_{d_i}}(\tilde{x}_i).$$

The mapping function of the ConvAC given in Eq. (8.1), h_y can be arranged into a scalar product $< \mathcal{A}^y, \mathcal{F} >$ between two tensors of order N and dimensions M.

Based on the expression of A^y presented in Eq. (8.5), this tensor accepts a CPD decomposition, as explained in Sec. 4.9. We aim then to estimate weights of the shallow network by decomposing A^y based on Eq.(8.1), thus training the network in an original way and avoiding the classical GD training and its convergence limitations [10].

Methodology

Let K be the number of training samples $X_1 \ldots X_k$, Z the number of rank-1 terms (CPD rank, equal to the number of feature maps generated by the conv filters) and M^N the number of entries in $\mathcal{A}^y$. Our method consists in:

1. Estimating $\mathcal{A}^y$ in a vectorized form $\mathcal{A}^y_{\text{vec}}$ through multilinear regression from Eq. 8.1.
2. Re-tensorizing $\mathcal{A}^y_{\text{vec}}$ into $\mathcal{A}^y$.
3. Factorizing $\mathcal{A}^y$ according to its CPD as in Eq. 8.5:

$$\mathcal{A}^{y} = \sum_{z=1}^{Z} a^{y}_{z} \mathbf{a}^{z,1} \otimes \cdots \otimes \mathbf{a}^{z,N},\ \mathbf{a}^{z,i} \in \mathbb{R}^{M_i}.$$

4. Feeding the ConvAC architecture with the optimal weights (W_{TD}) obtained from the tensor decomposition $W_{TD} = (\mathbf{a}^{z,i}, a^{y}_{z})_{z,i}$.
   We opt for coefficient sharing: $\forall i \in [N], \mathbf{a}^{z,i} = \mathbf{a}^{z}$ so that $\mathcal{A}^y$ becomes supersymmetric.

Experiments

First experiments are tunned on synthetic data for model validation. Two sets of normal random distribution images with means (μ_1, μ_2) and standard deviations (σ_1, σ_2), respectively, are synthesized. Adapted multivariate Gaussian filters f_{θ_1} and f_{θ_2}, respectively parametrized by (μ_1, σ_1) and (μ_2, σ_2) are applied as ConvAC representation functions. The hidden layer of the ConvAC is fed with these transformed images. Due to the

matching between the representation functions and the images, the convolutional and dense filters are theoretically expected to be close to the canonical vectors. However, training with classical algorithms such as SGD, adaptive learning rate optimization algorithms, GD with momentum in different configurations fails to converge. The gradients of the loss function tend to vanish within few epochs even for a variety of initialization schemes of the network weights.

As an illustrative toy example, we synthesize small images of size 4×6 generated from two equally-balanced normal random distributions $D_1 \sim \mathcal{N}(\mu_1 = 1, \sigma_1^2 = 0.5^2)$ and $D_2 \sim \mathcal{N}(\mu_2 = 3, \sigma_2^2 = 0.5^2)$. Images are partitioned into non overlapping patches of size 2×3. The Gaussian filters f_{θ_1} and f_{θ_2} are respectively parametrized $(\mu_{\theta_1} = 1, \sigma_{\theta_2}^2 = 1)$ and $(\mu_{\theta_2} = 3, \sigma_{\theta_2}^2 = 1)$. This setting is illustrated in Fig. 8.3. The variance of the representation function is increased to guarantee that more instances generated from D_i are captured by f_{θ_i} for $i = 1, 2$. The dataset measures 100 images, 50 from each class and it is partitioned into 75 images for training and 25 for validation. The network is trained for 100 epochs with SGD without momentum with learning rate $\lambda = 0.05$ and mini-batches of 25. The linear convolutional filters are 1×1 without overlapping. The binary cross entropy loss, introduced in Sec. 4.4.5, is used as optimization criterion and the categorical accuracy indicates the correctness of classification of the network. The net-

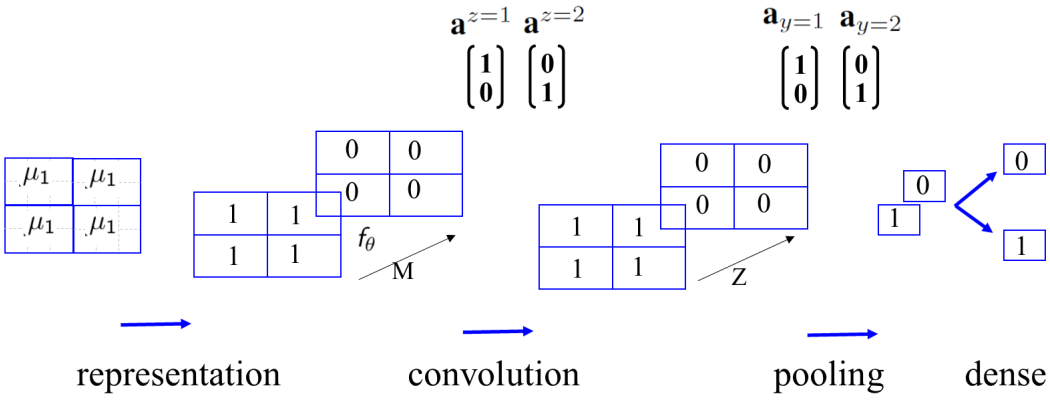

Figure 8.3: Classification of synthetic random images with ConvAC: Ilustrative toy example.

work succeeds in converging when properly initialized, but it fails when Gaussian noise of zero mean and a growing standard deviation is added to this initialization, as illustrated in Figs. 8.4-8.5, calculated for 100 MC simulations. Another problem related the floating point precision arises. Indeed, h_y needs to be normalized and positive in order to define a proper cross-entropy loss function. For this purpose we try to add a fully-connected layer to the output of the ConvAC, with a softmax activation function and weights frozen to the canonical vectors, thus transforming the score functions into

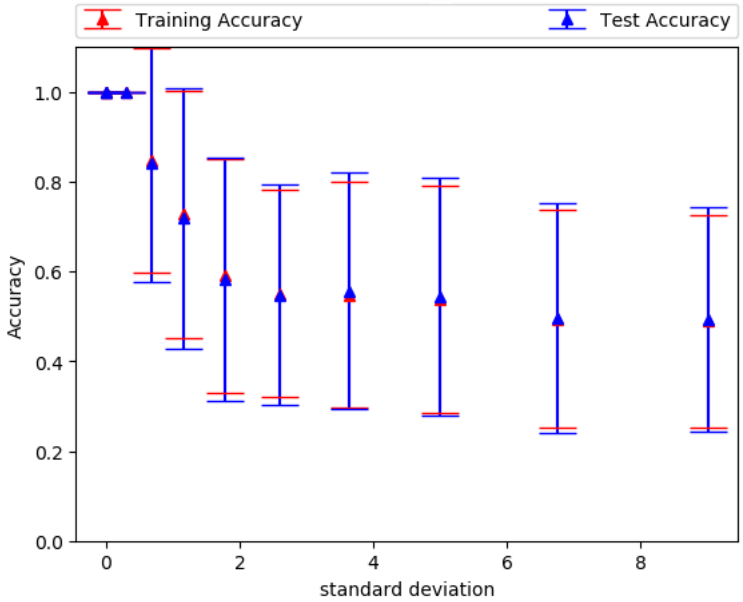

Figure 8.4: Effect of initialization on the convergence of ConvAC. Gaussian noise with standard deviation σ is added to the theoretical optimal weights of all trainable layers which are used as initialization. As σ grows the accuracy diminishes for both training and test samples.

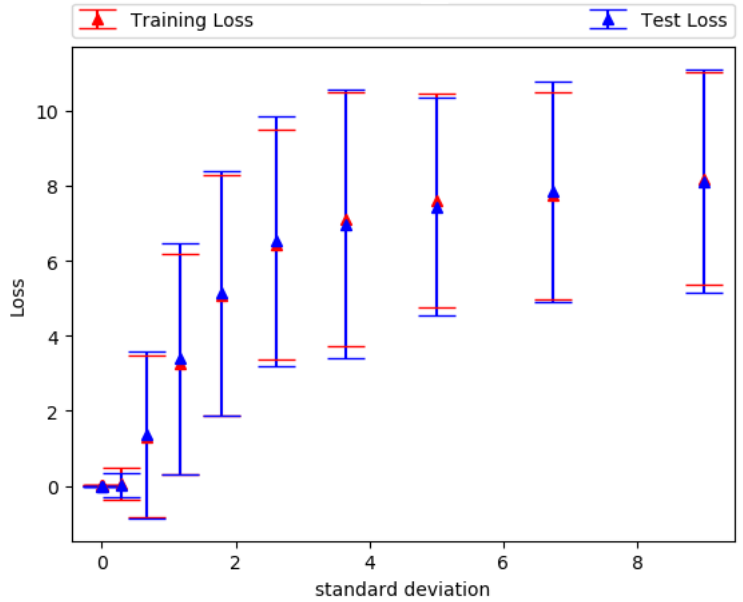

Figure 8.5: Effect of initialization on the convergence of ConvAC. Gaussian noise with standard deviation σ is added to the theoretical optimal weights of all trainable layers which are used as initialization. As σ grows the loss keeps growing for both training and test samples.

probabilities. But the resulting probabilities are all nearly equal to 0.5 since the response of h_y is low. CPD applied to $\mathcal{A}$ results in high residual error (near 1) because $\mathcal{A}$ is sparse. As a result the weights W_{TD} cannot be considered as an optimal algebraic initialization. Facing all these limitations, we abandon the idea and move to other network architectures to classify the dispersion EGMs.

Implementation: Experiments for coupled-symmetric CPD are conducted with MATLAB software using the TensorLab toolbox [113]. We use a MATLAB 8.4.0 version. Experiments of ML learning are conducted within Docker environment, running on Quadro P6000 GPU and Ubuntu 19.10 operating system using Python 3.6.9 programming language, Tensorflow 2.1.0 deep learning library for GPU and Keras API.

8.4 Classification Results

The identification of STD locations is performed with several ML techniques including TL. For this purpose, we benchmark the different EGM classification models detailed in Sec. 8.3. Their performance on test sets are given in Tab. 8.1. The values are calculated as the average over 5-fold CV [92] [93]. All standard deviations are inferior to 10^{-2}. Hence, all trained models have a low generalization error. Among the 35563 data points, there exist only 1804 samples labeled STD compared to 33759 non-STD samples. Regarding this low CIR (5%), we opt for ROS as proposed in 6. Results are commented next.

1. **Matrix Classification**: Both LeNet-STD with its three settings of f_{size} and MLR yield good and comparable performance. However, LeNet-STD slightly overperforms MLR with values of AUC and F1 equal to 92% and 53% respectively for MLR. The average values of AUC and F1 are 94% and 58% for LeNet-STD. The choice of f_{size} in LeNet-STD has a small impact on the classification performance but has an important effect on the computational cost. Indeed, F1 scores are 58%, 59% and 60%, respectively, with LeNet-STD$_{(4\times3)}$, LeNet-STD$_{(4\times200)}$ and LeNet-STD$_{(4\times300)}$ respectively. However, the training times of LeNet-STD$_{(4\times300)}$ and LeNet-STD$_{(4\times200)}$ are very high compared to LeNet-STD$_{(4\times3)}$ as shown in Tab. 8.2.

   Using ConvAC models would have been of a great interest regarding their theoretical generalization power. However preliminary experiments on synthetic data are not successful, as reported in Sec. 8.3.3. For this reason we decide to give up the idea of classifying EGM data (in matrix format) with ConvACs.

2. **2D Image Classification**: Second, we check if classifying 2D images would give good performance since this is the most intuitive input format. In practice, interventional cardiologist analyze visually EGM subplots to detect STD pattern, so

this format seems user-friendly. However, results in Tab. 8.1 show there is no significant enhancement in performance with 2D images compared to EGM matrices. Here again, we notice that LeNet-STD yields a better performance with an F1 score of 60% compared to 52% for MLR.

3. **3D Image Classification**: Applying TL of VGG16 model to our 3D images is performed as a proof of concept. VGG16-EGM yields good classification performance with values of Acc, AUC and F1 of 93%, 93% and 55%. This demonstrates that extracting features automatically from 3D plots with the use of conv blocks, trained on a different topic (natural images), does work. Besides the heavy computational cost of VGG16-EGM, as highlighted in Tab 8.2, synthesizing 3D images revealed costly too. Several extensions were benchmarked for storing and reading images in an optimal way. We recall that the study dataset includes more than 35500 samples and that storing 3D images would require 500 GB of disk space as JavaScript Object Notation (JSON) format for example.

4. **VAVp Classification**: The overall classification performance of VAVp time series is also good but less effective than the remaining data representations. Tab 8.1 shows that the performance of CNN-1D is slightly better than that of MLR. MLR beats significantly PCA-SVM, mainly in terms of F1 score. Moreover, the Gaussian kernel outperforms the linear one. The VAVp format does not yield the best value for any of the performance metrics considered in this evaluation and results in the lowest F1 scores, all below 0.46%.

8.4.1 Overall Assessment

Based on previous observations, we can claim that classifying raw EGMs stored in matrices with LeNet-STD$_{(4\times3)}$ yields a good balance between performance and computational cost. If we observe the values of precision and F1 across all experiments (Tab. 8.1), we notice poor values (around 30% with VAVp and inferior to 60% with the remaining features) even though values of TPR are high around 80% and 90% (except with PCA-SVM). However, values of PPV and F1 are high for the training set that is balanced with ROS. This low precision on the test sets can be explained by the fact that both test and validation sets are still highly imbalanced. So even though the trained models can identify better STD samples when the training set is augmented, as highlighted in 6, information about STD class is still less rich than that of non-STD due the redundancy caused by ROS. A low precision also indicates that the number of FP is important compared to the one of TP. This can be acceptable in biomedical data analysis as explained below. On the one hand, TPR is high which means that true STDs are being well de-

tected. On the other hand, it is always preferable to detect non-STDs as STDs than the inverse. This allows to tag these FP predictions during the mapping phase and give the cardiologist the opportunity to analyze them later, before ablation. Also, Tab. 8.2 proves that our solution can be deployed in real time since the prediction time (inference) of a new data sample can be performed in less than a second (maximal prediction time equal to 0.944 s for VGG16-EGM). If more efficient computational resources are deployed, prediction time can be further decreased.

Feature	**Model**	**Test classification performance**						
		Acc	**TPR**	**TNR**	**AUC**	**PPV**	**NPV**	**F1**
Matrix	MLR	0.924	0.902	0.847	0.929	0.389	0.991	0.533
	LeNet-STD$_{(4\times3)}$	0.936	0.911	0.862	0.94	0.436	0.992	0.579
	LeNet-STD$_{(4\times200)}$	0.937	**0.929**	0.881	0.940	0.443	**0.993**	0.589
	LeNet-STD$_{(4\times300)}$	**0.940**	0.928	0.879	**0.943**	**0.453**	0.99	**0.598**
2D Image	MLR	0.873	0.913	0.852	0.874	0.398	0.991	0.517
	LeNet-STD$_{(3\times3)}$	0.939	0.911	0.88	**0.943**	0.452	**0.993**	0.597
3D Image	VGG16-EGM	0.926	0.88	**0.928**	0.925	0.406	0.992	0.553
VAVp	PCA-SVM$_{\text{Lin}}$	0.791	0.670	0.535	0.805	0.128	0.970	0.207
	PCA-SVM$_{\text{Gaus}}$	0.837	0.732	0.616	0.849	0.179	0.976	0.277
	MLR	0.888	0.891	0.807	0.892	0.286	0.988	0.423
	CNN-1D	0.902	0.890	0.817	0.907	0.321	0.989	0.460

Table 8.1: Average classification performance on test set through 5-fold CV.

Time Complexity: PCA-SVM

$$O(\text{PCA}) = O(max(m,n)^2 \times min(m,n)) = O(m^2 \times n)$$
$$O(\text{SVM}_{\text{linear}}) = O(m \times r) \mid \text{worst case}$$

$$O(\text{SVM}_{\text{non-linear}}) \leq O(m^3 \times r)$$
$$O(PCA - \text{SVM}_{\text{linear}}) = O(m^2 \times n + m \times r) = O(m^2 \times n) \approx O(10^{12})$$

$$O(PCA - \text{SVM}_{\text{non-linear}}) = O(m^2 \times n + m^3 \times r) = O(m^3 \times r) \approx O(10^{15})$$

Feature	Model	Training cost		Prediction time (s)
		time	$\mathbf{nb_{tr\text{-}par}}$	
Matrix	MLR	00:01:49	400,672	0.139
	LeNet-STD$_{(4\times3)}$	00:11:04	1,289,458	0.643
	LeNet-STD$_{(4\times200)}$	00:24:57	1,214,322	0.625
	LeNet-STD$_{(4\times300)}$	00:37:51	1,175,922	0.616
2D Image	MLR	01:08:57	400,672	0.144
	LeNet-STD$_{(3\times3)}$	01:10:15	1,289,362	0.618
3D Image	VGG16-EGM	**14:35:28**	**100,795,778**	**0.944**
VAVp	PCA-SVM$_{\text{Lin}}$	00:10:14	$O(m^2 \times n)$	0.003
	PCA-SVM$_{\text{Gaus}}$	00:55:32	$O(m^3 \times r)$	**0.002**
	MLR	00:12:39	400,672	0.141
	CNN-1D	00:11:25	80,066	0.585

Table 8.2: Computational cost of training on balanced dataset and predicting the label of a test data sample. Time is given as hours(hh):minutes(mm):seconds(ss). Symbol $\text{nb}_{\text{tr-par}}$ stands for the number trainable parameters (weights) that might be inferior the total number of parameters in case of TL, when used as feature extractor. The number of variables in VAVp time series is $n = 2500$.

8.5 Summary and Conclusions

Automatic detection of atrial areas with STD pattern is a valuable decision-aid tool that would help interventional cardiologists identify potential ablation sites for treating persistent AF in a faster and more consistent way than current visual inspection. Despite their theoretical interest, the first attempts to train ConvAC architectures with a tensor decomposition in order to classify EGM recordings are unsuccessful. Indeed, the ConvAC architecture failed in classifying synthetic data. For this reason we give up this idea and move forward to more conventional classification algorithms. In order to identify STD EGMs, several features are extracted from EGM recordings whether automatically with TL and end-to-end NN training or in a hand-crafted way as in the VAVp time series. These different data representations are classified with the use of adapted ML algorithms, giving rise to a variety of EGM classifications models. Moreover, we study the effect of some hyperparameters and model settings like the choice of receptive fields and kernels in SVM. Based on the analysis of both the classification performance on the test set and the computational cost of the different classification models, the best performance is achieved with LeNet-STD and $f_{\text{size}} = 4 \times 3$ for classifying raw EGM matrices. The average performance over 5-fold CV reaches 94% of accuracy and AUC and an F1-score of 60%. The low precision and F1 score can be explained by the lack of insufficient amount of STD samples even though the issue of class imbalance is alleviated with DA. On the other hand, VGG16-EGM demonstrates that extracting features with conv blocks of the VGG16 model, which has been trained on natural images, works as well as shallower architectures like LeNet-STD, but is very expensive computationally.

The next chapter will study the automatic identification of the ablated sites among the mapped ones while taking into account the STD labels.

Chapter 9

Automatic Identification of STD Ablation Sites

Contents

9.1 Introduction

The previous chapter presented a comparative study of different EGM classification models. The purpose was to design a real time and optimized solution for the automatic identification of STD EGMs in STD-based ablation of persistent AF. Experimental results showed that baseline models such as LeNet-STD architecture applied to raw multichannel EGM records achieves the best performance in terms of AUC (94%) and computational cost. However, not all points labeled as STD in the mapping phase are later ablated in the ablation phase. The present chapter proposes to design a decision-aid solution that helps interventional cardiologists identify potential ablation sites automatically using ML tools. An annotation model is designed to automatically label the dataset of EGM recordings acquired during the mapping phase into two classes: 'ablated' and 'non-ablated'. We classify these mapped sites into the same categories using two baseline classifiers, previously used as a proof of concept in our exploratory study detailed

in Chap. 6, consisting in MLR and LeNet-STD. A binary label identifying whether the mapped site contains STD pattern according to the interventional cardiologist is combined to raw EGMs as classifiers' input. This additional input can be assimilated to a prior probability in STD-guided ablation. The performance of the different models are then benchmarked. ML techniques allow to automatically identify ablated sites and guide cardiologists in patient-tailored catheter ablation procedures for treating persistent AF. The recent study [121], provides evidence and examples about the high impact of ML in automatizing the diagnosis and treatment of AF.

9.2 Annotation Model

The study dataset contains a cohort of 35563 10-channel EGM signals of length 2.5 s composed of 1804 samples labeled STD compared to 33759 non-STD samples, as described in Sec. 5.2. Circularity transformation is applied to each data sample in order to mimic the circularity of PentaRay branches, as proposed in Sec. 5.2.4.

As explained in Sec. 5.2.1, the ablation intervention is a two-fold process, consisting in:

1. Mapping the atria and tagging STD sites
2. Ablating AF drivers (atrial locations), some of which may not necessarily have been mapped in advance

An annotated dataset is then needed to train supervised ML classifiers to automatically detect the ablated mapped points among all the mapped ones (including STD and non-STD). The labels are binary and should inform if the mapped point (data sample) has been ablated or not. However, this information (labels) is not provided by the CARTO system because:

1. Mapping and ablation phases are decoupled.
2. Ablation sites are not restricted to the mapped points tagged as STD. Indeed, there might be STD positions that are not ablated probably because they are isolated and do not form a cluster or because of anatomical reasons. Besides, there might be non-STD positions that are ablated because they are close enough to STD clusters or for anatomical reasons.

Multichannel EGM data are only available for mapped points, since the PentaRay catheter can only be used for mapping but not for ablation, added to the fact that the application of RF energy during ablation shots strongly disturbs EGM recordings.

For this purpose, we design a labeling model based on the proximity between both mapped $\mathcal{P}$ and ablated $\mathcal{V}$ point sets. Due to the limited precision of the ablation catheter and the manual ablation process, the catheter might be slightly misplaced around the desired position (D). In practice, the catheter may slightly move around the desired position D during an ablation shot, also called a session. Also D might correspond or not to a mapped point P.

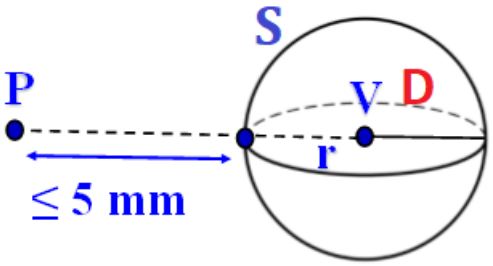

Figure 9.1: Sphere model of the ablation site. P represents a mapped point. V the centroid of the region visited by the ablation catheter, around point D, assumed to span a sphere $\mathcal{S}$. P is considered to have been ablated if the distance to the sphere surface is less than 5 mm.

Based on these observations, we propose to consider each area ablated around D as a sphere ($\mathcal{S}$) of center V and radius r. The Cartesian coordinates of V correspond to the average Cartesian coordinates of the points visited by the ablation catheter during the ablation session and r is the square root of the standard deviations of the corresponding coordinates. According to the partner interventional cardiologist, the distance between a mapped point P and the closest ablation zone $\mathcal{S}$ should be smaller or equal to 5 mm for P to be considered as ablated. This can be mathematically translated into the expression: $\|\overrightarrow{PV}\|_2 - r \leq 5$. Hence, for each ablation session, the mapped points that lie at most 5 mm away from $\mathcal{S}$ are attributed the 'ablated' label (P_{ab}), resulting in the set:

$$\mathcal{P}_{ab} = \{P_i \in \mathcal{P} : \exists V_j \in \mathcal{V}, d(P_i, V_j) \leq 5 + r_j\}, \tag{9.1}$$

where $d(P_i, V_j)$ represents the Euclidean distance between the points P_i and V_j. Mapped points not fulfilling Eq. (9.1) form the non-ablated class ($P_{\text{non-ab}}$). Based on this model, we obtain 15803 ablated and 19760 non-ablated points.

9.3 EGM Classification for the Identification of Ablation Sites

In order to identify potential ablation sites in STD-based ablation of persistent AF, we train baseline classifiers including MLR and LeNet-STD with $f_{\text{size}} = 4 \times 300$, already validated in a similar classification task in Chap. 6.

We conduct three types of experiments:

1. We train LeNet-STD and MLR with mini-batch GD and early stopping criteria. The receptive field of the first conv layer of LeNet-STD is chosen large enough to cover the information between all three contiguous leads along one AFCL and a half thus fully capturing potential STD patterns.

2. The CIR between labeled and non-labeled classes is equal to $\frac{\#P_{\text{ab}}}{\#P_{\text{non-ab}}} = 80\%$, where $\#A$ refers to the cardinality of the set A. As a result, classes are slightly imbalanced and the classifiers might face problems learning characteristics from the minority class (ablated) due to the insufficient number of P_{ab} samples.We propose to use a DA method called ROS that proved efficient in Chap. 6. The resulting models are denoted MLR-ROS and LeNet-ROS.

3. Based on the fact that both ablated and STD atrial areas should be spatially correlated, we concatenate the information about whether a data point is labeled STD or not with the previous classifiers' output. This additional input can be assimilated to a prior probability in STD-guided ablation. We build new architectures for MLR and LeNet-STD, denoted MLR-STDlab and LeNet-STDlab, by adding a sub-neural network (sub-NN) composed of a dense FC layer. The sub-NN gets the STD label as input and outputs a score function that will be concatenated with the output of the final decision layer of MLR and LeNet. We recall that for all the mapped points being classified, we have an STD label established by interventional cardiologists during the ablation intervention, as detailed in Chap. 5.

9.4 Classification Results

Five-fold CV is used to asses the generalization power of the trained classifiers. The classification results shown in Fig. 9.2 represent the average performance on the test dataset across the 5 folds of MLR (left) and LeNet (right) respectively. The metrics computed are the following, whose definitions are provided in Sec. 4.5:

- Accuracy (Acc).
- Area under the ROC curve (AUC).
- Sensitivity, also called true positive rate (TPR) or true rate of ablated samples.
- Specificity also called true negative rate (TNR) or true rate of non-ablated samples;
- precision, also called positive predictive value (PPV).

- Negative predictive value (NPV).
- F1 score.

The results show that the overall performance of LeNet-STD is better than that of MLR with values of Acc, AUC and F1-score respectively equal to 75%, 82%, 73% vs. 72%, 78%, 67%. ROS allowed us to increase the TPR from 76% to 84% while maintaining TNR at 70% which further increases the F1-score to 76% with LeNet-STD architecture. Concatenating with STD labels did not help improve neither decrease the classification performance. Also, the standard deviation of the classification metrics through CV does not exceed 10^{-2}. We conclude then that LeNet-STD-ROS is the best model for the identification of ablated points based on STD EGMs.

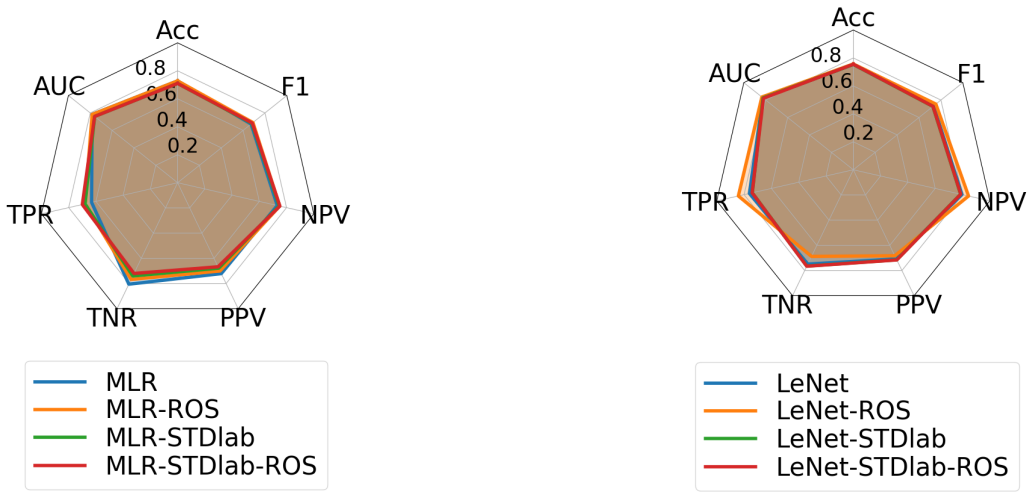

Figure 9.2: Classification performance on the test dataset. Values represent averages through 5-fold CV. **Left**: MLR classifier and variants. **Right**: LeNet-STD classifier and variants.

9.5 Summary and Conclusions

The automatic identification of potential ablations sites in STD-based ablation of persistent AF can help guide interventional cardiologists in patient-tailored procedures. For this task, we design a model to annotate the multichannel EGM dataset into ablated vs. non-ablated classes based on Euclidean distances between mapped and ablated points. ML tools are used to classify the labeled samples into the same categories. MLR and LeNet-STD are benchmarked whether combined with DA or not. Moreover, we concatenate the supervised classifiers with a sub-NN that takes as input the label of whether the considered data sample is tagged STD or non-STD. Five-fold CV is also used to asses

the robustness of trained models. Classification results on the test set show that LeNet-STD combined with ROS give the best performance with an F1 score of 76%. Aggregating STD label did not help improve the model's performance. This can be interpreted by the fact that the decision to ablate a patient takes into consideration more anatomical factors than the STD pattern. Moreover, the sub-NN used to analyze STD labels is shallow and might be not efficient.

In the next chapter, we will summarize the main objectives of the thesis and the challenges encountered. The most relevant results will be put forward. Then, we will discuss our methodology. Finally, we will give perspectives.

Part III

Conclusions

Chapter 10

Conclusions and Discussions

Contents

10.1 Summary of Contributions and Main Results

Among the existing treatments of AF, ablation based on the STD pattern of atrial EGMs remains a breakthrough therapy achieving promising procedural success rates. In practice, interventional cardiologists localize atrial sites displaying STD EGMs visually using the PentaRay multipolar mapping catheter. The present doctoral thesis aimed at designing an automatized, intelligent and real time decision-aid solution that helps interventional cardiologists in STD-guided ablation in persistent AF using adapted ML and DL tools. The proposed solution provides a three-fold contribution: first, we build a structured dataset of multichannel EGM recordings acquired by the PentaRay mapping catheter; second, we design a solution that identifies automatically STD EGMs while handling the highly imbalanced dataset issue; third, we propose to detect potential ablation sites in the atria automatically. Our work benefited from a continuous feedback from a partner cardiologist who approved the proposed solutions from a medical perspective. This aspect makes the results clinically relevant.

Overall, the proposed solution is among the first attempts at the application of sta-

tistical analysis and ML tools to automatically identify ablation areas in STD-based ablation. By providing interventional cardiologists with a real-time objective measure of STD, the proposed solution offers the potential to improve the efficiency and effectiveness of this fully patient-tailored catheter ablation approach for treating persistent AF.

The main results and contributions of this thesis can be summarized as follows:

10.1.1 Study Datasets

The experiments conducted in this work belong to two types of data: real and synthetic. The real dataset was exported from Nice CHU and contains a cohort of more than 35000 10-channel EGM recordings belonging to 16 different AF patients, acquired from STD-based ablation interventions and mapped with the PentaRay multipolar high-resolution catheter. The first experiments reported in Chaps. 6 and 7 studied only 23000 recordings from the LA mapping, because the LA harbors the main AF drivers and substrate. Then more data were exploited by adding EGMs acquired from the other cartographies in order to increase the size of the database. In addition to real EGMs, we proposed a realistic model of multichannel EGMs. The synthetic samples may contain parameterized STD and fractionation patterns, and proved useful in evaluating the impact of the VAVp in Chap. 7.

10.1.2 Automatic Identification of STD EGMs

The automatic identification of STD EGMs was addressed in Chaps. 6, 7 and 8 (see also [P1], [P2] and [P3]) with different approaches and tools. An exploratory study about the automatic classification of mapped sites into STD vs. non-STD groups was first performed with baseline classifiers. However, highly imbalanced dataset ratio hampered the classification performance. We tackled this issue by using adapted DA techniques that help achieve good classification. The overall performance was high with values of accuracy and AUC around 90%. Then, statistical measures and handcrafted features were extracted from the VAVp times series in order to characterize the STD pattern in EGMs. But no important dissimilarity was obtained between STD and non-STD categories in real data. Therefore, complementary experiments showed that the supervised classification of raw VAVp time series itself into the same categories is promising with values of accuracy, AUC, sensitivity and specificity around 90%. Finally, we benchmarked different EGM classification models in terms of performance and computational cost. Each model consists in a combination of a data representation and a classification algorithm. As a first attempt, we proposed to investigate shallow ConvAC architectures for their promising theoretical interest and their equivalence to CPD tensor decomposition but experimental results on synthetic data were unsuccessful. We moved forward

to more conventional supervised ML algorithms including transfer learning. Among the classification models considered, the best performance was achieved with a shallow CNN model for classifying raw EGM matrices. The average performance over cross-validation reached 94% of accuracy and AUC and an F1 score of 60%.

10.1.3 Automatic Identification of Ablation Sites

The automatic identification of ablation sites in STD-based intervention was addressed in Chap. 9 (see also [P4]). For this purpose, we labeled the points mapped during the first phase of the ablation intervention into ablated vs. non-ablated according to their proximity (Euclidean distance) to the ablated sites. Classification results on the test set showed that a shallow CNN gives the best performance with an F1 score of 76%. A complementary experience consisted in aggregating STD labels as a prior probability in the classification task, but this additional input did not help improve the model's performance.

10.2 Discussions and Limitations

The topic of automatizing the identification of ablation sites in STD-based ablation of AF from the analysis of multichannel EGM recordings is relatively recent and represents an unexplored area of research. We can then claim that our decision-aid solution is original and will constitute the state of the art for future studies. Therefore our solution is considered as a proof of concept about the applicability of ML, mainly CNN models, to the classification of EGM recordings acquired by the PentaRay catheter for STD-based ablation. The overall performance achieved with our proposed solution is good with values of accuracy and AUC around 90%, but better results are expected in a medical decision-aid tool. According to our partner cardiologist, an accuracy of 95% should be attained for the solution to be considered as sufficiently reliable for clinical use. Also, a higher number of patients would provide a richer EGMs database and increased variability across patients. Another limitation of our work consists in the fact that the real EGM dataset exported for the present study is not the product of a pure STD-based ablation but belongs to patients who were ablated with PVI, a conventional protocol, followed by the ablation of atrial sites displaying STD EGMs.

10.3 Perspectives

As mentioned in the discussion section, the present doctoral thesis provides a decision-aid solution for interventional cardiologists in STD-based ablation. Therefore, we can

consider the proposed solution as a proof of concept that the automatic classification of STD EGMs and the identification of potential ablation sites is feasible and can be performed with a high level of accuracy (more than 90%) while guaranteeing a low inference time (less than 1 s). The results are promising and raise some clinical and computational challenges as well as open questions.

Prominent perspectives of this PhD thesis include:

- Increasing the database to at least 100 persistent AF patients in order to provide more relevant clinical results and to meet the standards of a conventional clinical study [122].
- To enhance the convergence of the classifiers used in this thesis, further training algorithms, parameterizations and regularizations should be explored. In addition, investigating deeper and wider neural network architectures should be considered as larger networks can handle more efficiently the huge augmented dataset. For example, residual connections can be added to the conv blocks in deep CNN models such as in the study [98].
- Using LSTM models would be particularly interesting [63], as they are adapted to the classification of time series such as VAVp signals or even multichannel EGM signals (EGM matrix format). More details about LSTM are provided in Sec. 4.4.5.
- Interpreting, from a medical perspective, the features automatically synthesized by the deep architecture would bring knowledge to better characterize the persistent AF arrhythmia in general and the STD pattern in particular. For example, some studies invert the CNN models to visualize the features learned by the conv filters in face and object recognition tasks, as proposed in [43, Chap. 5].
- Predicting the short-term endpoints of ablation intervention whose definitions are provided in Sec. 3.3.1. Such an information is very important for the evaluation of the effectiveness of the STD-guided ablation protocol. However, such a study would require a higher number of subjects.
- Validating our results on further studies based on prospective data acquired from new persistent AF patients.
- After prospective validation of the proposed solution, develop real-time implementations and analyze their effectiveness to provide STD-based ablation in actual clinical practice.

References

[1] V. Zarzoso, O. Meste, P. Comon, D.G. Latcu and N. Saoudi. "Noninvasive cardiac signal analysis using data decomposition techniques," in P. Kornprobst, O. Faugeras, F. Cazals, Editors, *Modeling in Computational Biology and Biomedicine*, chap. 3, pp. 83–116, 2013.

[2] G. Lippi, F. Sanchis-Gomar and G. Cervellin . "Global epidemiology of atrial fibrillation: An increasing epidemic and public health challenge," *International Journal of Stroke*, 2020.

[3] C. T. January, L. S. Wann, J. S. Alpert, H. Calkins, J. E. Cigarroa, J. C. Cleveland, J. B. Conti, P. T. Ellinor, M. D. Ezekowitz, M. E. Field, et al. "2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: a report of the American College of Cardiology / American Heart Association Task Force on Practice Guidelines and the Heart Rhythm Society," *Journal of the American College of Cardiology*, 64(21), 2014.

[4] P. Kirchhof, S. Benussi, D. Kotecha, A. Ahlsson, D. Atar, B. Casadei, M. Castella, H. C. Diener, H. Heidbuchel, J. Hendriks, G. Hindricks, A. S. Manolis, J. Oldgren, B. A. Popescu, U. Schottenand B. Van Putte, P. Vardas, and ESC Scientific Document Group. "2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS," *European Heart Journal*, 37(38), pp. 2893–2962, 2016.

[5] C. T. January, L. S. Wann, H. Calkins, L. Y. Chen, J. E. Cigarroa, J. C. Cleveland, et al. "2019 AHA/ACC/HRS focused update of the 2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Rhythm Society," *Journal of the American College of Cardiology*, 74(1), pp. 104–132, 2019.

[6] J. G. Andrade, et al. “Cryoballoon or radiofrequency ablation for atrial fibrillation assessed by continuous monitoring: a randomized clinical trial,” *Circulation*, 140(22), pp. 1779–1788, 2019.

[7] K. Nademanee, J. McKenzie, E. Kosarand and et al. “A new approach for catheter ablation of atrial fibrillation: mapping of the electrophysiologic substrate,” *Journal of the American College of Cardiology*, 43(11), pp. 2044–2053, 2004.

[8] A. Verma, et al. “Approaches to catheter ablation for persistent atrial fibrillation,” *New England Journal of Medicine*, 372(19), pp. 1812–1822, 2015.

[9] J. Seitz, C. Bars, G. Théodore, S. Beurtheret, Nicolas N. Lellouche, M. Bremondy, A. Ferracci, J. Faure, G. Penaranda, M. Yamazaki, et al. “AF ablation guided by spatiotemporal electrogram dispersion without pulmonary vein isolation: a wholly patient-tailored approach,” *Journal of the American College of Cardiology*, 69(3), pp. 303–321, 2017.

[10] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*, MIT Press, 2016. www.deeplearningbook.org.

[11] J. Malmivuo and R. Plonsey. *Bioelectromagnetism: Principles and Applications of Bioelectric and Biomagnetic Fields*, Oxford University Press, USA, 1995.

[12] L. M. Davis, M. E. Rodefeld, K. Green, E. C. Beyer and J. E. Saffitz. “Gap junction protein phenotypes of the human heart and conduction system,” *Journal of Cardiovascular Electrophysiology*, 6(10), pp. 813–822, 1995.

[13] W. G. Stevenson and K. Soejima. “Recording techniques for clinical electrophysiology,” *Journal of Cardiovascular Electrophysiology*, 16(9), pp. 1017–1022, 2005.

[14] J. M. De Bakker. “Activation mapping: unipolar versus bipolar recording." *Cardiac Electrophysiology: From Cell to Bedside*, 2000, pp. 1068–1078, 1995.

[15] B. F Hoffman and P. F. Cranefield. “The physiological basis of cardiac arrhythmias,” *The American Journal of Medicine*, 37(5), pp. 670–684, 1964.

[16] B. M. Beckmann, P. Arne and S. Kääb. “Inherited cardiac arrhythmias: diagnosis, treatment, and prevention,” *Deutsches Ärzteblatt International*, 108(37), p. 623, 2011.

[17] J. Pellman and F. Sheikh. "Atrial fibrillation: mechanisms, therapeutics, and future directions," *Comprehensive Physiology*, 5(2), pp. 649–665, 2011.

[18] D. Dobrev. "Electrical remodeling in atrial fibrillation," *Herz*, 31(2), pp. 108–112, 2006.

[19] M. Haïssaguerre, M. Wright, M. Hocini and P. Jaïs. "The substrate maintaining persistent atrial fibrillation," *Circulation: Arrhythmia and Electrophysiology*, 1(1), pp. 2–5, 2008.

[20] J. A. Reiffel. "Drug choices in the treatment of atrial fibrillation," *The American Journal of Cardiology*, 85(10), pp. 12–19, 2000.

[21] J. B. Shea and W. H. Maisel. "Cardioversion," *Circulation*, 106(22), pp. 176–178, 2002.

[22] J. L. Cox, et al. "The surgical treatment of atrial fibrillation: III. Development of a definitive surgical procedure," *The Journal of Thoracic and Cardiovascular Surgery*, 101(4), pp. 569–583, 1991.

[23] P. A. Laizzo. "Catheter Ablation of Cardiac Arrhythmias," in *Handbook of Cardiac Anatomy, Physiology and Devices*. Springer Science and Business Media, chap. 25, pp. 356–364, 2009.

[24] S. S. Bun, T. Delassi, D. G. Latcu, M. El Jamili, A. Ayari, et al. "A comparison between multipolar mapping and conventional mapping of atrial tachycardias in the context of atrial fibrillation ablation," *Archives of Cardiovascular Diseases*, 111(1), pp. 33–40, 2018.

[25] M. Andronache, D. Nikola and V. Graziana. "High-resolution mapping in patients with persistent AF," *Arrhythmia and Electrophysiology Review*, 8(2), p. 111, 2019.

[26] E. Anter, C. M. Tschabrunn, F. M. Contreras-Valdes, J. Li and M. E. Josephson. "Pulmonary vein isolation using the Rhythmia mapping system: Verification of intracardiac signals using the Orion mini-basket catheter," *Heart Rhythm*, 12(9), pp. 1927–1934, 2015.

[27] J. D. Hummel. "Atrial mapping with basket catheters: a basket case?," *Journal of the American College of Cardiology, Clinical Electrophysiology*, 2(1), 2016.

[28] T. E. Walters and J. M. Kalman. "Human persistent atrial fibrillation is maintained by rotors: the jury is still out," *Circulation: Arrhythmia and Electrophysiology*, 8(3), 2015.

[29] D. Álvarez, F. Alonso-Atienza, J.L. Rojo-Álvarez, A. García-Alberola and M. Moscoso. "Shape reconstruction of cardiacischemia from non-contact intracardiac recordings: A model study," *Mathematical and Computer Modelling*, 55(5-6), pp. 1770–1781, 2012.

[30] "Products Catalogue 2020," *Biosense Webster*, 2020.

[31] Z. F. Issa, J. M. Miller and D. P. Zipes. "Focal Atrial Tachycardia," in *Clinical arrhythmology and electrophysiology: a companion to Braunwald's heart disease*. Elsevier Health Sciences, chap. 11, 2009.

[32] P. Teixeira, P. S. Cunha, A. S. Delgado, R. Pimenta, M. M. Oliveira and R. C. Ferreira. "PentaRay catheter in persistent atrial fibrillation ablation," *Revista Portuguesa de Cardiologia*, pp. 121–123, 2016.

[33] K. Nademanee, E. Lockwood, N. Oketani and B. Gidney. "Catheter ablation of atrial fibrillation guided by complex fractionated atrial electrogram mapping of atrial fibrillation substrate," *Journal of Cardiology*, 55(1), pp. 1–12, 2010.

[34] R. Parameswaran, A. M. Al-Kaisey and M. K. Jonathan. "Catheter ablation for atrial fibrillation: current indications and evolving technologies," *Nature Reviews Cardiology*, pp. 1–16, 2020.

[35] M. Haissaguerre, P. Jaïs, D. C. Shah, A. Takahashi, M. Hocini, et al. "Spontaneous initiation of atrial fibrillation by ectopic beats originating in the pulmonary veins," *New England Journal of Medicine*, 339(10), pp. 659–666, 1998.

[36] M. D. O'Neill, P. Jaïs, M. Hocini, F. Sacher, G. J. Klein, J. Clémenty, and M. Haïssaguerre. "Catheter ablation for atrial fibrillation," *Circulation*, 116(13), pp. 1515-1523, 2007.

[37] R. J. Hunter, I. Diab, M. Tayebjee, L. Richmond, S. Sporton, M. J. Earley, and R. J. Schilling. "Characterization of fractionated atrial electrograms critical for maintenance of atrial fibrillation: a randomized, controlled trial of ablation strategies (the

CFAE AF trial)," *Circulation: Arrhythmia and Electrophysiology*, 4(5), pp. 622–629, 2011.

[38] M. D. O'Neill, P. Jaïs, Y. Takahashi, A. Jönsson, F. Sacher, et al. "The stepwise ablation approach for chronic atrial fibrillation—evidence for a cumulative effect," *Journal of Interventional Cardiac Electrophysiology*, 16(3), pp. 153–167, 2006.

[39] M. Haïssaguerre, M. Hocini, P. Sanders, F. Sacher, M. Rotter, et al. "Catheter ablation of long-lasting persistent atrial fibrillation: Clinical outcome and mechanisms of subsequent arrhythmias," *Journal of Cardiovascular Electrophysiology*, 16(11), pp. 1138–1147, 2005.

[40] N. Oketani, J. Seitz, Ml. Salazar, A. Pisapia, J. Kalifa, J. J. Smit and K. Nademanee. "Ablation of complex fractionated electrograms is useful for catheter ablation of persistent atrial fibrillation: Protagonist point of view," *Heart Rhythm*, 13(10), pp. 2098–2100, 2016.

[41] B. Liu, A. Nuñez-Garcia, C. Tran and M. Wu. "Evaluating the Effectiveness of Atrial Fibrillation Ablation Guided by Spatiotemporal Electrogram Dispersion," *Circulation: Arrhythmia and Electrophysiology*, 142(3), pp. A16884–A16884, 2020.

[42] S. Iravanian, and J. J. Langberg. "Spatiotemporal organization during ablation of persistent atrial fibrillation," *Heart Rhythm*, 12(9), pp. 1937-1944, 2015.

[43] C. Francois. *Deep Learning with Python*, Manning Press, 2017.

[44] Y. LeCun, B. Yoshua, and H. Geoffrey. "Deep Learning," *Nature*, 521(7553), pp. 436–444, 2015.

[45] A.K. Jain, M.N. Murty and P.J. Flynn. "Data clustering: a review," *ACM Computing Surveys*, 31(3), pp. 264–323, 1999.

[46] R. Xu and D. Wunsch. *Clustering*, John Wiley and Sons, 10, 2008.

[47] D. Steinley. "K-means clustering: a half-century synthesis," *British Journal of Mathematical and Statistical Psychology*, 59(1), pp. 1–34, 2006.

[48] C. D. Manning, R. Prabhakar and S. Hinrich. *Introduction to Information Retrieval*, Cambridge University Press, pp. 378–401, 2008.

[49] S. H. Cha. "Comprehensive survey on distance/similarity measures between probability density functions," *City*, 1(2), p. 1, 2007.

[50] I. T. Jolliffe and C. Jorge, "Principal component analysis: a review and recent developments," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2065), pp. 2015–0202, 2016.

[51] M. E. Wall, A. Rechtsteiner and L. M. Rocha. "Singular value decomposition and principal component analysis," in *In A Practical Approach to Microarray Aata Analysis*, chap.5, Springer, pp. 91–109, 2003.

[52] A. J. Izenman. "Linear discriminant analysis," in *Modern Multivariate Statistical Techniques*, chap. 5, Springer, New York, pp. 237–280, 2013.

[53] D. I. Ellis and G. Royston. "Metabolic fingerprinting in disease diagnosis: biomedical applications of infrared and Raman spectroscopy," *Analyst*, 131(8), pp. 875–885, 2006.

[54] J. Ye, R. Janardan and Q. Li. "Two-dimensional linear discriminant analysis," *Advances in Neural Information Processing Systems*, 17, pp. 1569–1576, 2004.

[55] R. Martin-Clemente and V. Zarzoso. "LDA via L1-PCA of Whitened Data," in: *Proc. IEEE Transactions on Signal Processing*, 68, pp. 225–240, 2019.

[56] S. Dreiseitl and O. M. Lucila. "Logistic regression and artificial neural network classification models: a methodology review," *Journal of Biomedical Informatics*, 35(5-6), pp. 352–359, 2002.

[57] F. Melgani and B. Yakoub, "Classification of electrocardiogram signals with support vector machines and particle swarm optimization," *IEEE Trans. on Information Technology in Biomedicine*, 12(5), pp. 667–677, 2008.

[58] B. E. Boser, M. G. Isabelle and N. V. Vladimir, "A training algorithm for optimal margin classifiers," in *Proc. 5^{th} Annual Workshop on Computational Learning Theory*, USA, pp. 144–152, 1992.

[59] I. Steinwart and A. Christmann. *Support Vector Machines*. Springer Science and Business Media, 2008.

[60] K. Gurney. *An Introduction to Neural Networks*, CRC Press, 1997.

[61] A. Cichocki, U. Rolf and W. S. Roman. *Neural Networks for Optimization and Signal Processing*, 253, New York, 1993.

[62] T. M. Hagan, B. D. Howard and M. Beale. *Neural Network Design*. PWS Publishing Co., 1997.

[63] S. Hochreiter and J.Schmidhuber, "Long short-term memory," *Neural Computation*, 9(8), pp. 1735–1780, 1997.

[64] P. Baldi. "Autoencoders, unsupervised learning, and deep architectures," *Proceedings of ICML Workshop on Unsupervised and Transfer Learning*, 2012.

[65] S. Haykin and B. Widrow. *Least-Mean-Square Adaptive Filters*, John Wiley and Sons, 31, 2003.

[66] D.P. Kingma and J. Ba. "Adam: A method for stochastic optimization," *arXiv preprint*, arXiv:1412.6980, 2014.

[67] I. Goodfellow, et al. "Generative adversarial networks," *ArXiv Preprint*, arXiv:1406.2661, 2014.

[68] N. Srivastava, et al. "Dropout: a simple way to prevent neural networks from overfitting," *The Journal of Machine Learning Research*, 15(1), pp. 1929–1958, 2014.

[69] J. Schmidhuber. "Deep learning in neural networks: An overview," *Neural Networks*, 61, pp. 85–117, 2015.

[70] B. Pyakillya, N. Kazachenko and N. Mikhailovsky. "Deep learning for ECG classification," *Journal of Physics: Conference Series*, 913(1), IOP Publishing, 2017.

[71] A.Y. Hannun, P. Rajpurkar, M. Haghpanahi, et al., "Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network," *Nature Medicine*, 25(1), p. 65, 2019.

[72] U. R. Acharya, et al. "A deep convolutional neural network model to classify heartbeats," *Computers in Biology and Medicine*, 89, pp. 389–396, 2017.

[73] A. Krizhevsky and H. G. Hinton. "Object classification experiments," in *Learning Multiple Layers of Features From Tiny Images*, Master's thesis, Department of Computer Science, University of Toronto, chap. 3, 2009.

[74] N. Vinod and G. E. Hinton. "Rectified linear units improve restricted boltzmann machines," *International Conference on Machine Learning, ICML*, 2010.

[75] B. Xu, N. Wang, T. Chen and M. Li. "Empirical evaluation of rectified activations in convolutional network," *arXiv preprint*, arXiv:1505.00853 ,2015.

[76] C. Y. Lee, P. W. Gallagher, and Z. Tu. "Generalizing pooling functions in convolutional neural networks: Mixed, gated, and tree," *Artificial Intelligence and Statistics*. 2016.

[77] T. Williams and R. Li. "Wavelet pooling for convolutional neural networks," International Conference on Learning Representations, 2018.

[78] I. Strumberger, E. Tuba, N. Bacanin, M. Zivkovic, M. Beko and M. Tuba. "Designing convolutional neural network architecture by the firefly algorithm," in: *Proc. International Young Engineers Forum (YEF-ECE)*, Costa da Caparica, Portugal, pp. 59–65, 2019.

[79] Y. LeCun, L. Bottou, Y. Bengio and et al., "Gradient-based learning applied to document recognition," in: *Proc. of the IEEE*, 86(11), pp. 2278–2324, 1998.

[80] Alom, Md Zahangir, et al. "The history began from AlexNet: A comprehensive survey on deep learning approaches," *arXiv preprint*, arXiv:1803.01164 , 2018.

[81] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *International Conference on Learning Representations*, CA, 2015.

[82] J. Deng, et al. "ImageNet: A large-scale hierarchical image database," *IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, 2009.

[83] H. Poon and P. Domingo. "Sum-product networks: A new deep architecture," *IEEE International Conference on Computer Vision Workshops*, IEEE. pp. 689–690, 2011.

[84] N. Cohen, S. Or and S. Amnon. "On the expressive power of deep learning: A tensor analysis," *Conference on Learning Theory*. 2016.

[85] N. Cohen and A. Shashua. "Simnets: A generalization of convolutional network," arXiv preprint arXiv:1410.0781, 2014.

[86] F. L. Hitchcock. "The expression of a tensor or a polyadic as a sum of products," *Journal of Mathematics and Physics*. 6, pp. 164–189, 1927.

[87] W. Hackbusch and S. Kühn. "A New Scheme for the Tensor Representation," *Journal of Fourier Analysis and Applications*, 15(5), pp. 706–722, 2009.

[88] S. J. Pan and Q. Yang. "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering*, 22(10), pp. 1345–1359, 2009.

[89] L. Torrey and J. Shavlik. *Handbook of Research on Machine Learning Aapplications and Trends: Algorithms, Methods and Techniques*, IGI global, pp. 242–264, 2010.

[90] K. Weiss, T. M. Khoshgoftaar and D. Wang. "A survey of transfer learning," *Journal of Big Data*, 3(1), p. 9, 2016.

[91] K. Hajian-Tilaki, "Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation," *Caspian Journal of Internal Medicine*, 4(2), p. 627, 2013.

[92] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence*, 14(2), 1995.

[93] M. Kuhn and J. Kjell, *Applied Predictive Modeling*, Springer, 26, p. 70, 2013.

[94] J. G. Szekely and C. R. Rao. "Pre-limit and post-limit theorems for statistics," *Statistics Textbook and Monographs*, 161, pp. 411–422, 2000.

[95] D. N. Joanes and C. A. Gill. "Comparing measures of sample skewness and kurtosis," *Journal of the Royal Statistical Society: Series D (The Statistician)*, 47(1), pp. 183–189, 1998.

[96] M. Agnieszka and M. Grochowski, "Data augmentation for improving deep learning in image classification problem," *International Interdisciplinary PhD Workshop*, IEEE, 2018.

[97] C. Shorten and L. K. Taghi, "A survey on image data augmentation for deep learning," *Journal of Big Data*, 6(1), p. 60, 2019.

[98] A.Y. Hannun, P. Rajpurkar, M. Haghpanahi and et al. "Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network," *Nature Medicine*, 25, pp. 65–69, 2019.

[99] R.A. Harshman. "Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multimodal factor analysis," *UCLA Working Papers in Phonetics*, 16(1), p. 84, 1970.

[100] J. D. Carroll and J. Chang. "Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition," *Psychometrika*, 35, pp. 283–319, 1970

[101] L. De Lathauwer. *Signal Processing Based on Multilinear Algebra*, PhD thesis, Katholieke Universiteit Leuven, 1997.

[102] P. Comon. "Tensors: A brief introduction," *IEEE Signal Processing Magazine*, 31(3), pp. 44–53, 2014.

[103] T. G. Kolda and B. W. Bader. "Tensor decompositions and applications," *SIAM review*, 51(3), pp. 455–500, 2009.

[104] P. M. de Oliveira. *Tensor Modeling of the ECG for Persistent Atrial Fibrillation Analysis*, PhD Thesis, Université Côte d'Azur, 2019.

[105] V. Sandfort, K. Yan, P.J. Pickhardt and R.M. Summers. "Data augmentation using generative adversarial networks (CycleGAN) to improve generalizability in CT segmentation tasks," *Scientific reports*, 9(1), pp. 1–9, 2019.

[106] N. Nonaka and J. Seita, "Data Augmentation for Electrocardiogram Classification with Deep Neural Network," *ArXiv Preprint*, arXiv:2009.04398, 2020.

[107] P. Cao, X. Li, K. Mao, F. Lu, G. Ning, L. Fang and Q. Pan. "A novel data augmentation method to enhance deep neural networks for detection of atrial fibrillation," *Biomedical Signal Processing and Control*, 56, p.101675, 2020.

[108] A. Goldberger, et al. "PhysioBank, PhysioToolkit, and PhysioNet: Components

of a new research resource for complex physiologic signals," *Circulation [Online]*. 101(23), pp. 215–220, 2000.

[109] G. D. Clifford, et al. "AF classification from a short single lead ECG recording: The PhysioNet/computing in cardiology challenge 2017," *Computing in Cardiology (CinC)*, Rennes, pp. 1–4, 2017.

[110] J. L. Wells, R.B. Karp, N. T. Kouchoukos, W. MacLean, T. N. James, A. L. Waldo. "Characterization of atrial fibrillation in man: studies following open heart surgery," *Pacing Clin Electrophysiol*, 1, pp. 426–438, 1978.

[111] K. T. Konings, et al. "High-density mapping of electrically induced atrial fibrillation in humans," *Circulation*, 89(4), pp. 1665–1680, 1994.

[112] J. Yang, M.N. Nguyen, P.P. San, X. Li and S. Krishnaswamy, "Deep convolutional neural networks on multichannel time series for human activity recognition," *in 24th International Joint Conference on Artificial Intelligence*, 2015.

[113] N. Vervliet, O. Debals, L. Sorber, M. Van Barel and L. De Lathauwer. *Tensorlab 3.0*, URL: `http://www.tensorlab.net`, 2016.

[114] M. Kachuee, S. Fazeli and M. Sarrafzadeh, "ECG Heartbeat Classification: A Deep ransferable Representation," in *Proc. IEEE International Conference on Healthcare Informatics (ICHI)*, New York, NY, 2018, pp. 443–444, 2018.

[115] Y. Qihang, et al. "Multi-class arrhythmia detection from 12-lead varied-length ECG using attention-based time-incremental convolutional neural network." *Information Fusion*, 53, pp. 174–182, 2020.

[116] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, et al. "Attention is all you need," *ArXiv Preprint*, arXiv:1706.03762, 2017.

[117] M. Alfaras, M. C. Soriano, and S. Ortín. "A fast machine learning model for ECG-based heartbeat classification and arrhythmia detection," *Frontiers in Physics*, 7, 2019.

[118] G. B. Moody and R. G. Mark. "The impact of the MIT-BIH Arrhythmia Database," in: *Proc. IEEE Engineering in Medicine and Biology*, 20(3), pp. 45–50, 2001).

[119] M. Rodrigo, et al. "Classification of Atrial Fibrillation by Deep Learning of Electrogram Shapes versus Rate and Regularity," *Circulation*, 142(Suppl 3), pp. A14738–A14738, 2020.

[120] S. I. Duque, A. Orozco-Duque, V. Kremen, D. Novak, C. Tobón and J. Bustamante. "Feature subset selection and classification of intracardiac electrograms during atrial fibrillation," *Biomedical Signal Processing and Control*, 38, pp. 182–190, 2017.

[121] K. C. Siontis, X. Yao, J. P. Pirruccello, A. A. Philippakis and P. A. Noseworthy. "How will machine learning inform the clinical care of atrial fibrillation?," *Circulation*, 127(1), pp. 155–169, 2020.

[122] National Institute Of Aging, "What are clinical trials and studies?," `https://www.nia.nih.gov/health/what-are-clinical-trials-and-studies`, reviewed in 2020.

Publications Derived From this Work

[P1] A. Ghrissi, D. Almonfrey, R. Almeida, F. Squara, V. Zarzoso, J. Montagnat. "Data augmentation for automatic identification of spatiotemporal dispersion electrograms in atrial fibrillation ablation using machine learning," in *Proc. 42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2020*, Montreal, CA 2020.
Selected as Europe geographic finalist in the student paper contest.

[P2] A. Ghrissi, F. Squara, V. Zarzoso and J. Montagnat. "Identification of spatiotemporal dispersion electrograms in persistent atrial fibrillation ablation using maximal voltage absolute values," in: *Proc. 28th European Signal Processing Conference*, Amsterdam, Netherlands, 2021.

[P3] A. Ghrissi, D. Almonfrey, F. Squara, V. Zarzoso, J. Montagnat. "Identification of Spatiotemporal Dispersion Electrograms in Atrial Fibrillation Ablation Using Machine Learning: A Comparative Study," submitted, January, 2021.

[P4] A. Ghrissi, F. Squara, V. Zarzoso and J. Montagnat. "Identification of ablation sites in persistent atrial fibrillation based on spatiotemporal dispersion of electrograms using machine learning," in: *Proc. Computing in Cardiology Conference*, Rimini, Italy, 2021.

[P5] M. Sallem, A. Ghrissi, A. Saadaoui, V. Zarzoso, "Detection of cardiac arrhythmias from varied length multichannel electrocardiogram recordings using deep convolutional neural networks," *PhysioNet/CinC Challenge*, in: *Proc. Computing in Cardiology*, Rimini, Italy, 2020.