



HAL
open science

Deep probabilistic generative model for the inverse problem of electrocardiography

Tania Bacoyannis

► **To cite this version:**

Tania Bacoyannis. Deep probabilistic generative model for the inverse problem of electrocardiography. Medical Imaging. Université Côte d'Azur, 2021. English. NNT : 2021COAZ4111 . tel-03621337

HAL Id: tel-03621337

<https://theses.hal.science/tel-03621337>

Submitted on 28 Mar 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE DE DOCTORAT

Modèle génératif probabiliste pour le problème inverse
de l'électrocardiographie

Tania-Marina BACCOYANNIS

INRIA, Équipe EPIONE

LIRYC, Équipe MULTIMODAL DATA SCIENCE

Thèse dirigée par Maxime SERMESANT

Soutenue le 17 Décembre 2021

Présentée en vue de l'obtention du grade de DOCTEUR EN AUTOMATIQUE, TRAITEMENT
DU SIGNAL ET DES IMAGES de l'UNIVERSITÉ CÔTE D'AZUR.

Devant le jury composé de :

Oscar CAMARA

Universitat Pompeu Fabra

Rapporteur

Yves COUDIÈRE

Université de Bordeaux

Rapporteur

Vicente ZARZOSO

Université Côte d'Azur

Président

Hubert COCHET

CHU Bordeaux, Liryc

Co-encadrant

Maxime SERMESANT

Inria Sophia Antipolis, Liryc

Directeur de thèse

Membres du Jury

Titre français

Modèle génératif probabiliste pour le problème inverse de l'électrocardiographie

Titre anglais

Deep Probabilistic Generative Model for the Inverse Problem of Electrocardiography

Devant le jury composé de :

Président du jury

Vicente ZARZOSO, Professeur - HDR, Université Côte d'Azur

Rapporteurs

Oscar CAMARA, Professeur, Universitat Pompeu Fabra

Yves COUDIÈRE, Professeur - HDR, Université de Bordeaux

Examineurs

Oscar CAMARA, Professeur, Universitat Pompeu Fabra

Yves COUDIÈRE, Professeur - HDR, Université de Bordeaux

Hubert COCHET, Professeur, CHU Bordeaux, Liryc

Maxime SERMESANT, Directeur de recherche - HDR, Inria Sophia Antipolis, Liryc

Abstract

Electrocardiographic imaging (ECGI) is an exceptional resource in cardiology practice and research, owing to its non-invasive manner to study cardiac electrical activity abnormalities by means of acquiring ECGs via multi-electrodes placed on the torso. This enthralling approach is largely based on solving an inverse ill-posed problem of ECGI. However, to date, no method for solving this problem has been sufficiently convincing in order to establish ECGI as a clinical modality of first choice. Thus, the aim of this thesis has been to endeavor to significantly contribute towards the advancement of ECGI implementation, in order to establish it in the near future as a preferentially routine method of arrhythmia diagnostic in the clinics. The work undertaken in this thesis has allowed us to propose a novel formulation of ECGI problem based on deep learning. Specifically, the core of the model we developed is a variational and conditional autoencoder using deep generative neural networks. Our unique model is able to produce advanced maps of cardiac activation by conditioning the electrical activity of the heart with signals measured on the torso surface and the cardiac geometry.

The thesis manuscript follows the three-step evolution of our deep learning-based model. In the first step, we treated the problem in a two-dimension space. We reformulated the entire ECGI problem employing a probabilistic generative model, which is a variational and conditional autoencoder based on convolutional neural networks. The model was trained and validated on a synthetic database that we generated from CT images of a patient. Using the eikonal model, we simulated cardiac activation maps considering single onsets. To compute the corresponding Body Surface Potentials (BSPs), we modelled the torso electrodes by a 10×10 Cartesian grid. Considering the torso domain homogeneous and infinite, we used the dipole formulation with the Mitchell-Schaeffer model to generate the BSP. Secondly, the 2D model was further generalised to a 3D model, which was trained on a simulated database from the CT images of five patients and then evaluated on a new cardiac geometry. The last part of this thesis work consists in evaluating our proposed 3D method on realistic clinical data. The clinical data processing includes: acquisition and segmentation of CT images, along with the measurement and analysis of non-invasive ECG signals on the torso as well as of invasive catheter-based mapping of activation times and unipolar/bipolar voltages via a conventional electro-anatomical system (for validation purposes). In this part, we simulated cardiac activation maps considering multiple onsets. The torso surface potentials are generated on two Cartesian grids of 10×13 electrodes each. Moreover, here the activation maps are no longer

conditioned by the binary mask of cardiac image but rather by speed maps, although they are still conditioned by the potentials on the thoracic surface. The model was trained on five patients anatomies and tested on two additional patients clinical data.

Our probabilistic generative model is superior to classic methods in several aspects. In particular, thanks to convolutional neural networks, our model is able to learn the spatio-temporal interactions between the signals, whereas most existing methods for ECGI either solve each time step independently or use a temporal prior without spatial correlations. We should also highlight that the correlation between the imaging substrate and the signals is also learned, allowing us to integrate any 3D image information into the ECGI-based model, which is difficult to achieve in the classical formulation. To our knowledge, our ECGI method is the first method able to reconstruct activation patterns in the whole myocardium volumetrically and at high spatial resolution. Our novel framework could be rapidly translated into routine clinical workflow for non-invasive, accurate diagnostic and ECGI-based testing in large cohorts of arrhythmia patients.

Keywords: Cardiac activation; Computational Modelling; Data Processing; Deep Learning; Electrocardiography; Generative Model; Inverse Problem.

Résumé

L'imagerie électrocardiographique (ECGI) est une ressource exceptionnelle dans la pratique et la recherche en cardiologie, en raison de sa manière non invasive d'étudier les anomalies de l'activité électrique cardiaque par l'acquisition d'ECG via des électrodes multiples placées sur le torse. Cette approche fascinante est largement basée sur la résolution du problème inverse de l'ECGI. A ce jour, encore aucune méthode pour résoudre ce problème connu comme étant mal posé n'a su convaincre suffisamment pour rendre l'ECGI la modalité de prédilection en clinique. Ainsi, l'objectif de cette thèse a été de contribuer de manière significative à l'avancement de l'ECGI, afin de l'établir dans un avenir proche comme méthode préférentielle en routine clinique pour le diagnostic des arythmies. Le travail entrepris nous a permis de proposer une nouvelle formulation du problème ECGI basée sur le Deep Learning. Le cœur du modèle développé est un autoencodeur variationnel et conditionnel utilisant des réseaux neuronaux génératifs profonds. Notre modèle unique est capable de générer des cartes d'activation cardiaque en conditionnant l'activité électrique du cœur avec des signaux mesurés à la surface du torse et la géométrie cardiaque.

Le manuscrit de la thèse suit l'évolution en trois étapes de notre modèle basé sur le Deep Learning. Premièrement, nous avons traité le problème dans un espace à deux dimensions. Nous avons reformulé l'ensemble du problème de l'ECGI en employant un modèle génératif probabiliste, qui est un autoencodeur variationnel et conditionnel basé sur des réseaux de neurones convolutifs. Le modèle a été entraîné et validé sur une base de données synthétique que nous avons générée à partir des images CT d'un patient. Nous avons simulé les cartes d'activation en utilisant le modèle eikonal en considérant des points d'initialisation uniques. Pour calculer les potentiels à la surface du torse correspondants, nous avons modélisé les électrodes du torse par une grille cartésienne de 10×10 . Considérant que le domaine du torse est homogène et infini, nous avons utilisé la formulation du dipôle couplée au modèle de Mitchell-Schaeffer pour générer les potentiels de la surface thoracique. Deuxièmement, le modèle 2D a été généralisé en un modèle 3D, qui a été entraîné sur une base de données simulées à partir des images CT de cinq patients, puis évalué sur une nouvelle géométrie cardiaque. La dernière partie de ce travail de thèse consiste à évaluer la méthode 3D que nous avons proposée sur des données cliniques réelles. A des fins de validation, le traitement des données cliniques comprend: l'acquisition et la segmentation d'images CT, l'analyse de signaux ECG non invasifs sur le torse ainsi que la cartographie invasive par cathéter des temps d'activation

et des tensions bipolaires via un système électro-anatomique conventionnel. Dans cette partie, nous avons simulé les cartes d'activation cardiaque et les potentiels à la surface du torse en appliquant le modèle eikonal sur deux grilles cartésiennes de 10×13 électrodes chacune. De plus, ici les cartes d'activation ne sont plus conditionnées par le masque binaire de l'image cardiaque mais par des cartes de conductivité, bien qu'elles soient toujours conditionnées par les potentiels de la surface thoracique.

Notre modèle génératif probabiliste est supérieur aux méthodes classiques sur plusieurs aspects. En particulier, grâce aux réseaux de neurones convolutifs, notre modèle est capable d'apprendre les interactions spatio-temporelles entre les signaux, alors que la plupart des méthodes existantes résolvent chaque pas de temps indépendamment ou utilisent un a priori temporel sans corrélations spatiales. De plus, la corrélation entre le substrat d'imagerie et les signaux est également apprise, ce qui nous permet d'intégrer toute information d'image 3D dans le modèle basé sur l'ECGI, ce qui est difficile à réaliser dans la formulation classique. À notre connaissance, notre méthode ECGI est la première méthode capable de reconstruire le modèle d'activation dans l'ensemble du myocarde de manière volumétrique et à haute résolution spatiale. Notre nouveau framework peut être rapidement intégré dans le workflow de routine clinique pour effectuer des tests diagnostiques non invasifs et précis basés sur l'ECGI dans de larges cohortes de patients souffrant d'arythmie.

Mots-clés: Activation cardiaque; Apprentissage profond; Electrocardiographie; Modélisation numérique; Modèle génératif; Problème inverse; Traitement de données.

Acknowledgements

First of all, I would like to thank my thesis director Maxime Sermesant. I would like to thank you Maxime for your benevolent supervision, the trust you have given me first of all by offering me the chance to join the team for my thesis but also during all these years. I would also like to thank you for your natural generosity in sharing your knowledge, your patience, your support, your encouragement and your communicative enthusiasm to work, to move forward, to persevere and whatever the difficulties encountered or to be overcome, to always believe and remain positive. Thank you for all this as well as for all the opportunities you gave me to collaborate but also to learn by following the training courses that interested me and this until Moscow.

Thanks to Nicholas Ayache, Hervé, Irene, Marco and Xavier, for making the wonderful Epione team so dynamic and ambitious.

I would like to thank Professor Hubert Cochet, my thesis co-supervisor, for his help on the clinical aspects but also for his encouragement, his support and his enthusiasm for my project.

I would also like to thank the members of my jury, Professors Oscar Camara and Yves Coudière for having read my manuscript with such rigour, for your comments and remarks as well as the feedback on my research work. Finally, I would like to thank Professor Vicente Zarzoso for having accepted to chair my thesis jury.

I would also like to thank Professor Josselin Duchateau for the valuable clinical data he shared with me. It certainly gave me a lot of trouble, but it allowed me to take my project to the next level, by evaluating it on clinical data.

I would also like to thank Nadjia Kachenoura and Professor Demetri Psaltis for their support, presence and all they gave me. But also Professor Vladimir Dotsenko for all the hours spent together solving physics or mathematics problems at the blackboard during my university years. Thanks to Dominique Lavault for everything.

More broadly, I would like to thank all the members of the Epione team and the Inria Sophia Antipolis centre for the good working atmosphere and all the good times spent together. Thank you Isabelle for your kindness and all that you did to facilitate my administrative tasks. I would like to express my gratitude to Buntheng and Mihaela for

their generosity, advice and support. A big thank you to Nicolas Cedilnik, Mehdi Juhoor and Pauline Migerditichan without whom my thesis would have been different. Pawel, Clément, Qiao thank you for your advice and listening, Yingyu for your good cakes and your kindness, Josquin, Raphaël, Jairo, Marta and all the others and to the best cat-sitters ever: Luigi and Jaume.

I would like to thank all my relatives and friends here in Paris, in Greece or in the four corners of the globe, and to all those who have contributed, in one way or another, to the achievement of this project.

And finally, I could not conclude these few lines without thanking my mama and my brothers for their unconditional support, their love and their constant encouragement, which increase every day the desire and the strength I have to move forward. This thesis is dedicated to you.

Thanks to life!

Financial Support

The research leading to these results has received European funding from the ERC starting grant ECSTATIC (715093) and French funding from the National Research Agency grant IHU LIRYC (ANR-10-IAHU-04). The project was also supported by the Inria Sophia Antipolis - Méditerranée, "NEF" computation cluster and by UCA Doctoral School EDSTIC.



Contents

1	Clinical and Methodological context and background	3
1.1	Introduction	3
1.1.1	Motivation	3
1.1.2	Manuscript Organisation	5
1.2	Clinical context and practice	6
1.2.1	From cell to electrical activity of the heart	6
1.2.1.1	Cardiac Anatomy	6
1.2.1.2	Cardiac Electrophysiology	7
1.2.2	Clinical modalities for Cardiac Electrophysiology Mapping	10
1.2.2.1	Electrocardiogram	10
1.2.2.2	Intra-cardiac electrocardiograms	11
1.2.2.3	Body Surface Potentials Mapping	12
1.2.2.4	Computed tomography scan	13
1.2.2.5	Magnetic resonance imaging	13
1.3	Methodological Context and State-of-the-art	14
1.3.1	Cardiac Electrophysiological models	14
1.3.2	Electrocardiographic Imaging	15
1.3.2.1	Forward problem	15
1.3.2.2	Inverse problem	16
1.3.3	Brief review of Deep Learning models and Related work	17
1.3.3.1	Introduction to Deep Learning models	17
1.3.3.2	Convolutional neural networks	19
1.3.3.3	Autoencoder	20
1.3.3.4	Variational autoencoder	21
1.3.3.5	Deep probabilistic generative model applied to ECGI	21
1.3.3.6	State-of-the-art of Deep Learning models for ECGI	22
1.4	Conclusion	23
2	Evaluation in 2D on Synthetic Data	25
2.1	Introduction	26
2.2	Context	27
2.2.1	Electrocardiographic Imaging	27
2.2.1.1	Forward Problem	27
2.2.1.2	Inverse Problem	27

2.2.2	Deep Learning	27
2.2.2.1	Variational Autoencoders	27
2.2.2.2	Conditional Variational Autoencoders	29
2.2.2.3	β -Variational Autoencoders	29
2.3	Methods	30
2.3.1	ECGI Forward Problem: Data Simulation	30
2.3.2	Models Structure and Training	30
2.3.2.1	VAE for BSP	30
2.3.2.2	VAE for Activation Maps	31
2.3.2.3	β -CVAE for Activation Maps from BSP and Images	31
2.4	Results	32
2.4.1	Evaluation of Body Surface Potentials VAE	32
2.4.2	Evaluation of Activation Maps VAE	33
2.4.3	Evaluation of β -CVAE	33
2.5	Discussion & Conclusion	34
3	Generalisation to 3D and Unseen Cardiac Geometries	35
3.1	Introduction	36
3.2	Methods	39
3.2.1	Deep Learning Approach	39
3.2.2	Supervised Learning Loss Function	42
3.2.3	Synthetic Data Generation: ECGI Forward Problem	43
3.2.4	Computational Details	45
3.3	Results	46
3.4	Discussion	47
3.5	Conclusion	48
4	Fusion of Body Surface Potential Mapping and Cardiac Imaging for ECGI Evaluation	49
4.1	Introduction	49
4.1.1	Clinical database	49
4.1.2	Motivation	50
4.1.3	Current data status and brief overview of the strategies employed	50
4.1.4	Material	54
4.2	Clinical Data processing	54
4.2.1	Extract the data	54
4.2.2	Segmentation	54
4.2.3	Cardiac registration	56
4.2.4	Generic Torso mesh	57
4.2.5	Cardiac mesh	61
4.2.6	Transfer the invasive measurements to the optimal mesh	65
4.2.6.1	Invasive data processing	65

4.2.6.2	Data Projection	68
4.2.6.3	Assess projection methods	69
4.2.6.4	New cardiac mesh with real Bipolar local activation times	78
4.3	Conclusion	80
5	Evaluation of Deep Learning ECGI formulation with Clinical Data	81
5.1	Introduction	83
5.2	Materials and Methods	85
5.2.1	Clinical Data	85
5.2.2	Data Optimisation	87
5.2.3	Data sets Construction	88
5.2.3.1	Test data set: transfer from clinics to anatomical models	88
5.2.3.2	Data Generation: training/validation data sets	90
5.2.4	Deep generative model for ECGI	91
5.2.5	Computational Details	93
5.3	Results	94
5.4	Discussion	99
5.5	Conclusion	101
6	Conclusions	103
6.1	Contributions of the Thesis	103
6.2	Perspectives	105
6.2.1	From 12-lead ECG	106
6.2.2	Interpretability of the latent space	107
6.2.3	A collaborative framework for ECGI clinical data processing . . .	107
	Bibliography	109

Glossary:

AI: Artificial Intelligence.

BEM: Boundary Element Method.

BSP: Body Surface Potentials.

BSPM: Body Surface Potentials Mapping.

CNN: Convolutional Neural Network.

CT scan: Computed Tomography scan.

CVAE: Conditional Variational Autoencoder.

CVD: Cardiovascular Diseases.

DL: Deep Learning.

EAM: Electro-Anatomical Mapping.

ECG: Electrocardiogram.

ECGI: Electrocardiographic Imaging.

EGMs: Intra-Cardiac Electrocardiograms.

EKG: Electrocardiography.

FEM: Finite Element Method.

FMM: Fast Marching Method.

KL: Kullback-Leibler divergence.

KPCA: Kernel Principal Component Analysis.

LV: Left Ventricle.

ML: Machine Learning.

NN: Neural Network.

RV: Right Ventricle.

SA: Sinoatrial Node.

TMP: Transmembrane Potential.

VAE: Variational Autoencoder.

VI: Variational Inference.

WHO: World Health Organisation.

Clinical and Methodological context and background

Contents

1.1	Introduction	3
1.1.1	Motivation	3
1.1.2	Manuscript Organisation	5
1.2	Clinical context and practice	6
1.2.1	From cell to electrical activity of the heart	6
1.2.2	Clinical modalities for Cardiac Electrophysiology Mapping	10
1.3	Methodological Context and State-of-the-art	14
1.3.1	Cardiac Electrophysiological models	14
1.3.2	Electrocardiographic Imaging	15
1.3.3	Brief review of Deep Learning models and Related work	17
1.4	Conclusion	23

The human heart, because of its complexity and central location in the body, has always been a subject of fascination and debate due to the related it attributed functions and properties. The properties associated with the heart are functional, symbolic and even poetic. Although contradictory according to civilisations, times and beliefs, there is one point everyone has always agreed on: the heart plays an essential and vital role in our lives.

1.1 Introduction

1.1.1 Motivation

While 1.8 million COVID-19 deaths were reported in 2020, every year cardiovascular diseases (CVD) cause about 17.9 million deaths, representing 32% of all global deaths, according to the World Health Organisation (WHO)¹. This number is rising steadily as the world population ages. The WHO estimates that by 2030, 23.6 million people will die from CVD.

¹https://www.who.int/health-topics/cardiovascular-diseases#tab=tab_1

The heart is a complex organ acting as an intricate electromechanical system. Indeed, a periodic electrical impulse generated by the heart itself and propagated throughout it allows for a rhythmic contraction and blood pumping. The pathologies associated with cardiac abnormal electrical activity constitute a major source of morbidity/mortality, especially those caused by structural heart disease.

In order to be able to localise the origin of a cardiac conduction abnormality generating aberrant rhythms, the cardiologist needs a representation of the electrical potential on the heart surface. To establish diagnosis, prognostic and therapeutic strategies, clinicians routinely rely on electrocardiograms (ECG) and medical imaging. Although these are the two most commonly used investigative modalities and are both non-invasive, they do not provide an essential and intramural cardiac activation map.

The therapy of choice to eliminate the foci of certain conduction disorders, is to burn out these sources using energy-based radio frequency (RF) thermal ablation. Prior to performing the ablation, the catheter is introduced into the heart chamber to record electrical signals (i.e., electrocardiograms, EGMs) and to identify arrhythmia foci based on EGMs' morphology, typically using electro-anatomical mapping (EAM) and navigation systems. This invasive technique is time-consuming (e.g. hours) and does not always localise intramural defective areas (i.e., deep in the myocardium), nor does completely reconstruct the voltage maps on the epicardial surface. As a consequence, the success rate of this potentially curative therapy is limited to only 60%.

The device used in predictive therapy, the implantable defibrillator (ICD), often fails to detect all true arrhythmic events and to always prevent sudden cardiac death.

While the spectrum of therapies that can be categorised as preventive and curative are being improved, the detection of foci by more accurate and less invasive methods still needs to be properly addressed.

As an alternative to invasive EAM and simplistic ECGs, a recent technique named Electrocardiographic Imaging (ECGI) was developed using non-invasive multi-modal data. This ECGI method allows the reconstruction and visualisation of cardiac electrical information such as body surface potentials. In particular, ECGI requires: 1) anatomical information of the heart and torso from 3D imaging; 2) body surface potentials (BSP) measured with a vest of multi-electrodes (e.g. 64-256); 3) the localisation of these electrodes on the torso surface; and, 4) mathematical methods to solve an inverse problem (which is ill-posed in Hadamard's sense). At the expense of a loss of accuracy, most approaches solve this problem by employing regularization techniques for the stability of the solutions.

To address the limitations associated with such approaches, one can employ artificial intelligence methods. For instance, Deep Learning (DL) has drastically improved the

state-of-the-art in many fields [LeCun, 2015] and in particular in medical image analysis [Litjens, 2017]. Furthermore, DL-based methods are not only very efficient for solving inverse problems in general [Arridge, 2019] but also precisely in medical imaging for image reconstruction, for example [Yedder, 2020], [Hammernik, 2018].

Developments in the fields of signal and image processing, biophysical and electrophysiological modelling as well as ECGI and DL (among others), could help clinicians increase the accuracy of diagnosis, improve the treatment of arrhythmia conditions, and thus try to reduce the devastating impact of such heart diseases.

This thesis specifically focuses on the development of a new DL-based approach to tackle the inverse problem of ECGI.

1.1.2 Manuscript Organisation

In this thesis, we approach the inverse problem of Electrocardiography as a learning problem. Being convinced that Deep Learning is not just a trend, but a powerful tool with many advantages, especially for solving inverse problems, we explored this perspective.

- To choose the appropriate Deep Learning model for this ECGI problem, we needed to: have a background in cardiac anatomy and electrophysiology; understand the measurements of the different modalities; study the different components, variants of neural networks, different learning techniques; and, review the state of the art of Deep Learning models in the context of ECGI. This ensemble, which is an essential prerequisite for a deep understanding of the problem we tried to solve, is covered in **Chapter 1** of this thesis.
- In **Chapter 2**, we have reformulated the whole problem of ECGI by a probabilistic generative model, which we believe is the most appropriate model for the problem. More precisely, this is a variational and conditional autoencoder based on convolutional neural networks. The choice of each element of the model is sensible, and so are the answers to the questions we asked ourselves before selecting this model. This chapter highlights the chronology of the implementation of our new formulation, namely we started by simply studying the variational autoencoders intended to reconstruct their input data. Finally, the conditional variational autoencoder (CVAE) developed is able to solve the inverse problem of ECGI in two dimensions (estimates slices of cardiac activation maps).
- In **Chapter 3**, the work conducted in **Chapter 2** is extended to the third dimension. It is currently possible to estimate volumetric cardiac activation maps for a new unseen cardiac geometry from the decoder part of the model after having trained the CVAE with synthetic data considering single pacing.

- **Chapter 4** is dedicated to the implementation of a clinical data processing pipeline from heterogeneous data. The objective was to fuse the data collected from multiple modalities (CT-scan, ECGI, EAM) to form a database of clinical data that can be used in practice to evaluate our model.
- **Chapter 5** consists of a modified version of CVAE which was presented in Chapter 3, in order to become better adapted to realistic clinical data, for evaluation purposes. The modifications are made both in the methods used to simulate the data as well as in the model itself. The model is evaluated on the clinical data obtained from the data fusion we have performed.
- Finally, in **Chapter 6**, we conclude by describing our main scientific contributions to this field and discussing potential future work and perspectives.

1.2 Clinical context and practice

1.2.1 From cell to electrical activity of the heart

As the subject of this thesis is not biology, only the electrophysiological elements essential to the overall understanding of cardiac function in the context of electrical activity will be presented.

The biological and electrochemical mechanisms involved in the electrical and mechanical activation of the heart are very complex. The subject of this thesis does not deal with this; therefore, only the main elements necessary for the global understanding of these phenomena will be described.

1.2.1.1 Cardiac Anatomy

Our body needs oxygen in order to function. The heart is the central driving force of the bloodstream, which provides this oxygenation to various tissues and organs. The heart is a muscle which, thanks to its internal structure, works like a mechanical pump and an electrical system. It is divided into four chambers, as shown in Figure 1.1: the left and the right chambers (left/right atria and left/right ventricles, respectively). The valves are structures that keep the blood flowing between atria and ventricles, while the septum is the interventricular muscle that separates the left ventricle (LV) from the right ventricle (RV).

The right chambers (RA/RV) receive the deoxygenated blood through the venous system, sending it to the lungs which will enrich it with oxygen. The left chambers (LA/LV)

receive this oxygen-rich blood and pump it into the rest of the body through the aorta and a network of arteries. This briefly presented process is thus driven primarily by the contractile function of the heart.

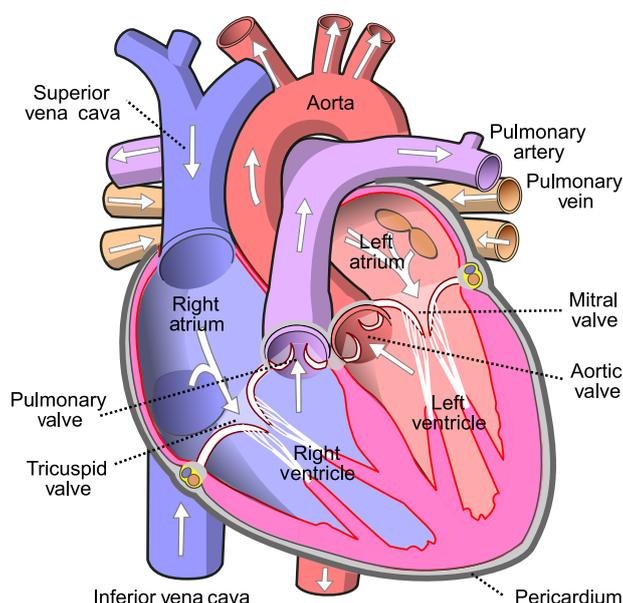


Fig. 1.1: Diagram of the main components of the human heart anatomy, showing the two parts of the heart synchronously pumping blood into different parts of the circulatory system. The blood flow direction is represented by the white arrow. The blue colour represents the deoxygenated blood being directed from the organs to the lungs. In pink is the oxygenated blood flowing from the lungs into the left ventricle, while in red is the blood full in oxygen ejected by the LV (which passes through the aorta before reaching the systemic circulation again). ; image from <https://commons.wikimedia.org/>.

Typically, one second is the time needed for a complete cardiac cycle of a healthy adult human heart. This cycle consists of two phases: systole and diastole. During systole the ventricles contract and pump out the blood, while during the diastole phase the ventricles relax and fill in with blood.

Structurally, the heart wall is composed of three tissue layers (see Figure 1.2a): the endocardium which is in direct contact with the blood; the myocardium which is the muscular tissue of the heart (i.e., responsible for the cardiac contraction); and, the outermost layer which is named epicardium.

1.2.1.2 Cardiac Electrophysiology

At the cellular level, the heart is composed of cells (named cardiomyocytes) that have their own distinct properties. For instance, the cells located in the sinus atrial node contribute to the initiation of the electrical impulse (excitation). The specialized Purkinje fibers allow a fast propagation of the impulse received from the atria to the ventricular

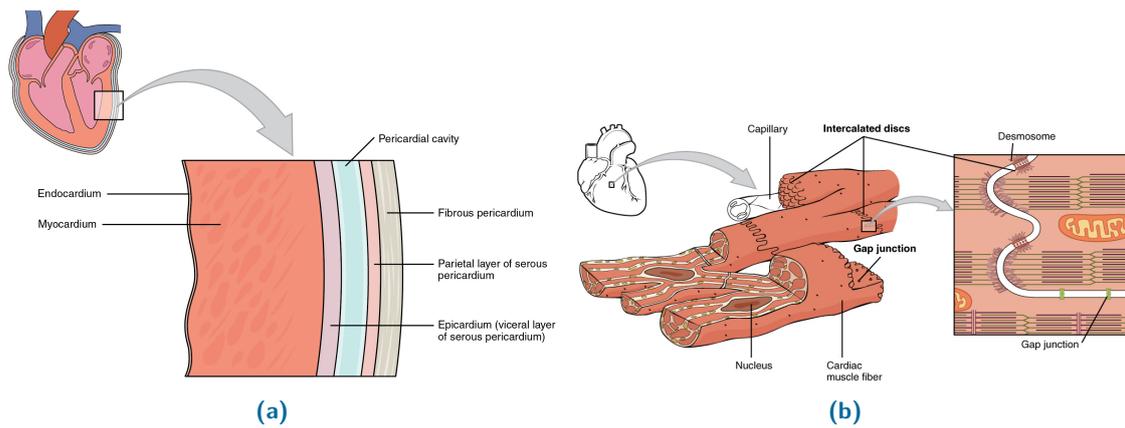


Fig. 1.2: (a) Membranes and layers of the heart wall, (b) Heart muscle composition and organization. From Wikimedia commons.

cells, while the myocardial cells (named cardiomyocytes) are responsible to respond to this excitation by contracting. Myocardial cells are organized in fibers and bundles, which which form sheets. Cardiomyocytes are connected through specialised structures: intercalated disks as shown in Figure 1.2b and desmosomes, forming a functional syncytium. The gap junctions (protein channels) enable the transmission of action potentials (i.e., the electrical impulse), while the desmosomes play a mechanical role allowing the muscle cells to be networked together through filaments' adhesion, and thus to act together during fibres' contraction.

Each myocardial cell is surrounded by a membrane that separates the extracellular medium from the intracellular medium. These two media are electrically conductive and have opposite charges (membrane potentials), being defined by the concentration of principal ions: Calcium (Ca^{2+}), Potassium (K^+) and Sodium (Na^+). The first study [HODGKIN, 1952] of ion channels was developed for the squid axon, and described the generation and propagation of action potential. In brief, before receiving any electrical impulse, the voltage potential is negative in the intracellular medium and positive outside the cell. The differences in ionic concentration induce a difference in potential, which is named transmembrane potential (TMP), and typically has a value of $-90mV$ [Zipes, 2018] in the resting state. The cell is said to be polarised. After a stimulus is applied, the gradient of ionic concentrations changes, and consequently the TMP is modified. Cellular activation is achieved by the depolarisation of the cells (i.e., a reversal of the potential in the extra- and intracellular media). Depolarisation is the first of the three main phases determining the evolution of the TMP, as shown in Figure 1.3a. During the depolarisation phase, the sodium channel (Na^+) is activated and the TMP reaches a peak of approximately $+20mV$. At the plateau phase, the TMP is quite balanced due to a slow deactivation of K^+ and Na^+ channels, whereas the calcium channel is open. Finally, the repolarisation occurs after the closing of Ca^{2+} channel and when the K^+

channel is reactivated. Then, the cardiomyocyte returns to its resting state and the TMP reaches again its resting potential of -90mV .

This variation in potential from the cellular activation to the resting state constitutes the action potential.

The sinoatrial node (SA), commonly referred to as our natural pacemaker, is responsible for the cardiac rhythm due to its ability to spontaneously generate electrical signals. Under normal conditions, the sinus rhythm originates in the SA node which is located in the right atrium, as illustrated in Figure 1.3b. The sinus rhythm of a healthy human heart at rest corresponds to a heart rate about 60 to 100 beats per minute. Next, the electrical signal propagates to the atrioventricular node located in the septum, then through the His bundle and into the network of Purkinje fibres. The cells of the Purkinje network are in charge of activating the cardiac myocytes in the ventricles.

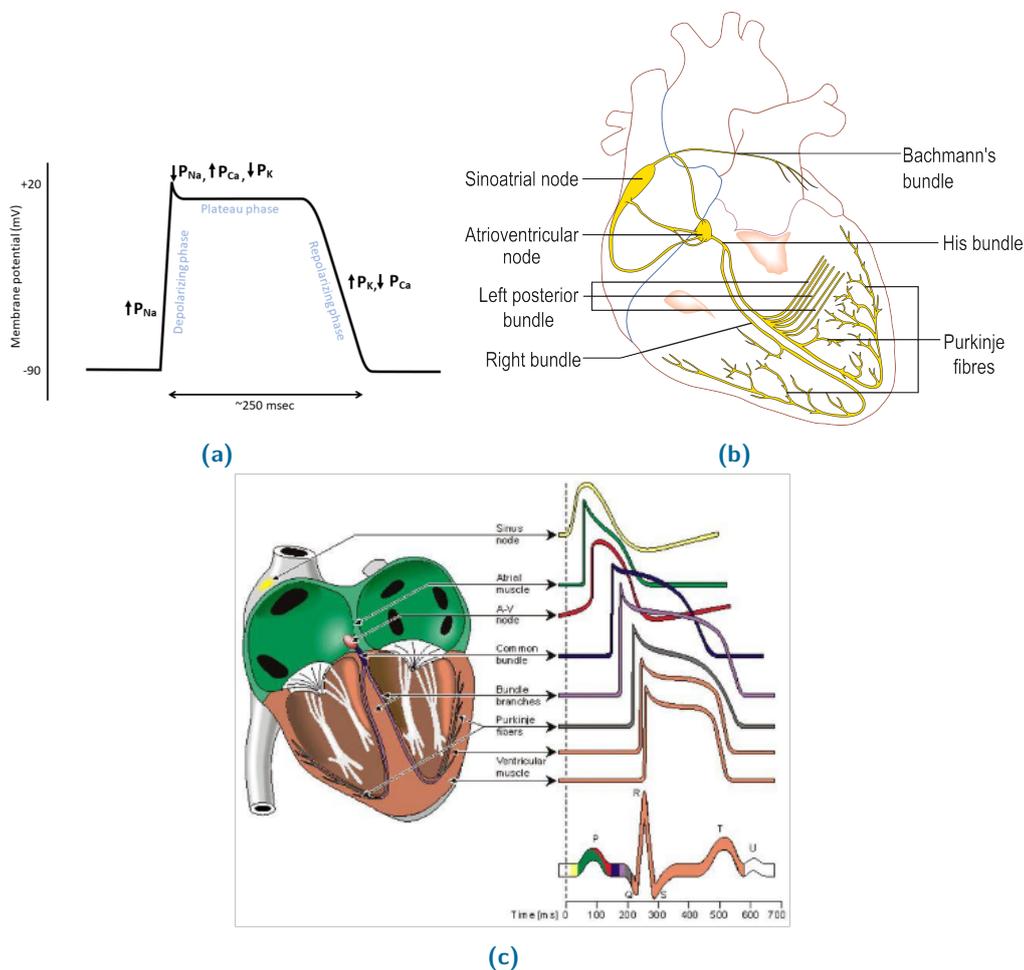


Fig. 1.3: (a) Action Potential response of a cardiac cell and ionic flow through the cell membrane at each stage. (b) Cardiac Electrical conduction system (in yellow). (c) Cardiac conduction system and action potentials. Representation of the corresponding Electrocardiogram. Images (a), (b) are taken from Wikimedia commons, and (c) from [Malmivuo, 1995].

The cardiac electrical activity groups together all the elementary electrical activities represented by each cellular action potentials, as shown in Figure 1.3c.

1.2.2 Clinical modalities for Cardiac Electrophysiology Mapping

1.2.2.1 Electrocardiogram

We have seen earlier that the electrical activity of the heart can be observed by measuring the extracellular potential directly on the surface of the heart and also from the potentials on the surface of the torso. However, the differences in potentials in the heart induce an electrical field that propagates from the heart to other organs. Thus, by measuring the differences in potentials on the surface of the skin, we can obtain information about the heart electrical activity. This non-invasive process can be recorded by an Electrocardiogram (EKG) or ECG. The electrocardiography method consists in placing several (3-12) electrodes on the surface of the torso, each having precise locations due to the fact that the cardiac electric field is space-dependent. These locations are standardised by The American Heart Association (AHA). At each moment in the cardiac cycle, the potential differences between two electrodes are measured and represented on the electrocardiogram (see Figure 1.4). The ECG signal amplitude along with the depolarization and repolarization times are used to characterise the ECG signal and define clinically important time intervals (as illustrated in Figure 1.4). Specifically, the P-wave corresponds to the atria depolarisation, while the ventricular depolarisation is expressed by the QRS complex and the ventricular repolarisation by the T-wave, respectively. The PR-segment corresponds to the propagation of the activation through the atrioventricular node and the Purkinje fibres, and QT-segment corresponds to the plateau of the ventricular action potential. The signal amplitude and the intervals' length are altered in the presence of pathologies such as electrical conduction anomalies.

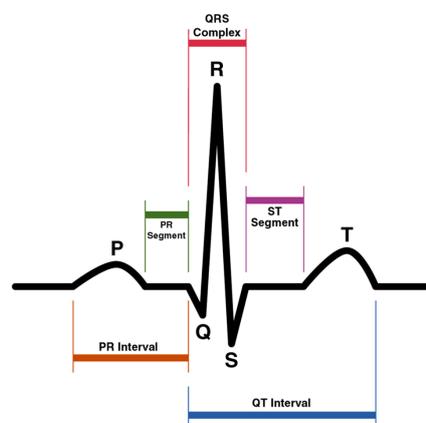


Fig. 1.4: Standard sinus rhythm electrocardiogram. From Wikimedia commons.

1.2.2.2 Intra-cardiac electrocardiograms

Electro-anatomical Mapping (EAM) is an invasive technique that allows the in-vivo study of cardiac electrophysiology, by providing intra-cardiac electrocardiograms (EGMs). EAM provides three-dimensional point-to-point mapping of intracardiac electrical activation. EGMs are recordings of the extracellular potential differences between two electrodes. The cardiac electrical activity is recorded locally in EGMs, as the electrodes are in direct contact with the myocardial tissue.

The cardiologist uses a mapping catheter which is inserted either through the femoral artery (for left heart study) or through the femoral vein (for right heart study), depending on which part of the heart will be investigated. Two different catheter-based approaches are used to record the signals: a contact-mapping approach and a non-contact approach, respectively. The former captures electrical signals by touching the endocardial (or epicardial) surface with the catheter, while the latter uses a catheter with a multi-electrode array over an inflatable balloon inside the heart chamber. The surrounding electrical activity is measured remotely.

Regardless of the mapping system chosen for the procedure, two types of EGMs are recorded during EAM: unipolar and bipolar [Alberts, 2002].

- Unipolar EGMs: these measurements consist of recording the potential variations between a positive electrode on the myocardial tissue and a "neutral" or "indifferent" electrode ([Ringwala, 2019; Stevenson, 2005]) placed at a distance from the heart, so that it is the weakest possible signal, or ideally null. This neutral electrode is placed either in the inferior vena cava or Wilson's central terminal. The electrode in the myocardial tissue is considered positive, such that when the wavefront moves towards it, the deflection in the EGM will be positive and vice versa. The maximum negative slope is reached when the wavefront is located below the recording electrode. These unipolar EGMs allow for the generation of activation maps and the identification of the initialisation point of the activation. However, it is sometimes difficult to separate the far field (due to depolarisation of tissue far from the recording site) from the local field because the far field has a higher amplitude than the local field and therefore it could overshadow the latter [Aliot, 2009].
- Bipolar EGMs: For bipolar EGMs, the two electrodes are placed in close proximity to each other. The EGMs are generated by calculating the difference in potential recorded between these two electrodes. The method of generating bipolar EGMs makes the detection of local activation more complex. In general, the absolute maximum amplitude of the EGM is used to identify the local activation. However, in the case of complex and fractionated bipolar EGMs [Gardner, 1985; Aliot, 2009;

Stevenson, 2005], this benchmark is less reliable than with unipolar electrograms. As the far-field signals in the two electrodes are nearly identical, bipolar EGMs are very useful for studying local signals.

One of the commercially available systems for the contact-mapping is the CARTO[®] system², shown in Figure 1.5. It uses a low energy electromagnetic field through three coils that are placed under the patient during the procedure. The combination of this field with the electric field of the catheter allows a precise 3D spatial localisation of the catheter.

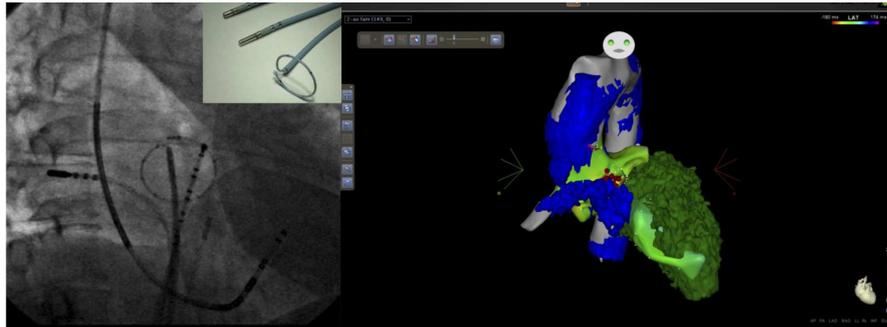


Fig. 1.5: CARTO[®] navigation system: with (Left) fluoroscopic images, (Centre) the CARTO[®] contact mapping catheters, (Right) the electroanatomical mapping. (Images courtesy of CHU, Bordeaux)

The clinical electrophysiological data available in this work was acquired using the contact-mapping approach, and will be presented in section 4.

Although EAM allows the reconstruction of the cardiac cavity to be mapped thanks to the localisation of the catheter, the entire myocardium cannot be reconstructed due firstly to the time required to image it, which is clinically unfeasible; secondly because of the approximations made on its anatomy; thirdly because of the errors on catheter's localisation induced by the respiratory motion as well as the cardiac movements. Furthermore, depending on the system used, the anatomical reconstruction algorithms are not the same and each of them has its own weaknesses. All of these limitations can hinder the integration of EAM information with the anatomical information of traditional imaging systems. In [Soto-Iglesias, 2016], the authors compare different fusion methods for interpolating activations recorded via EAM. A common solution to improve spatial sampling is to use multiple mapping catheters [Aliot, 2009].

1.2.2.3 Body Surface Potentials Mapping

The technology for Body Surface Potentials Mapping (BSPM) is very similar to the routine 12-lead ECG in that it uses electrodes placed on the surface of the torso. Previously,

²<https://www.jnjmedicaldevices.com/en-US/product-family/3d-navigation>

we saw that the position of the electrodes for ECG have been standardised. Numerous investigations have been carried out on the number of electrodes to be used and their ideal positions. In the context of solving the inverse ECGI problem, [Jiang, 2009] has shown that the position of the electrodes has a strong influence on the quality of the reconstruction of the cardiac electrical activity when a limited number of electrodes is used. BSPM uses up to 256 electrodes distributed over the anterior and posterior surfaces of the torso. The objective of this technique is to reconstruct the electrical activity on the surface of the myocardium based on these potentials measured on the surface of the torso. This brings the question of solving the inverse problem of the ECGI, which will be detailed in section 1.3.2.2. However, ECGI also requires anatomical data obtained by a cardiac imaging modality, which in our case is the CT-scan. Our ultimate goal will be to replace the invasive EAM approach with a non-invasive integrated ECGI-CT method.

1.2.2.4 Computed tomography scan

Computed tomography scan (CT scan) is an advanced medical imaging technique that is governed by the physical principle of X-ray absorption by various organs. An X-ray generator is placed in a scanner (ring) around the patient's body, while a receptor is placed diametrically opposite. In order to acquire an image of the target area in the patient, the X-ray source irradiates the patient. The radiation passes through the patient's tissue, and it is measured by the receptor in form of signals which is digitally converted into a CT image by a computer. Due to a radiation-matter interaction principle, the photons absorbed by a tissue depend on their attenuation coefficient. Each tissue absorbs (attenuates) differently the X rays; thus, the photons captured by the receptor are less numerous than the photons transmitted by the source. This technique provides images of 2D slices and further computer processing allows the 3D reconstruction and visualization of the selected volume. The images obtained are grey level images, where each pixel corresponds to an absorption (attenuation) coefficient of the tissue. In some cases, a iodine-based or barium-based contrast agent is injected into the patient, to increase the contrast between adjacent structures (with different perfusion characteristics), in order to better visualise certain regions of interest where a pathology is suspected.

1.2.2.5 Magnetic resonance imaging

Magnetic resonance imaging (MRI) is based on the magnetic relaxation processes of the proton, usually of the hydrogen atom in the water molecule. Its nucleus (consisting of a proton) has a spin which in the absence of a magnetic field has a random orientation, inducing a zero macroscopic magnetic moment. When the patient is placed in the MRI scanner, the spins create a magnetisation vector parallel with the magnetic field of the scanner. Following a sequence of excitation RF pulses, the received signal corresponds to the measurements of the current induced by spin's rotation back to its equilibrium posi-

tion. From this signal and complex mathematical calculations (using Fourier methods), one can obtain three-dimensional images with excellent contrast.

Usually, the image characteristics (such as spatial resolution, signal-to-noise ratio or contrast) along with the specific study requirements, are used to select one imaging modality over another. MRI is known to have better contrast than CT, while CT has a better spatial resolution than MR. The voxel size (which defines the spatial resolution) is very important for the study of heterogeneity in chronic healed infarct, since this is associated in late stages of the infarct with wall thinning of less than 1mm. Advances in the field of cardiac CT makes possible to study these heterogeneities using a spatial resolution of 1mm, whereas MRI cannot currently achieve resolutions $< 2\text{mm}$ in the clinics (in routine scans). Moreover, the contraindications for gadolinium-enhanced MR imaging and for individuals having various metallic implants, the high cost of MR equipment and its less-wide availability, must often be taken into account. In contrast, the CT scan is less expensive, requires less human intervention and a faster examination than MRI, and is also feasible for patients with pacemakers, cochlear implants, valves with a metallic component, etc.

1.3 Methodological Context and State-of-the-art

1.3.1 Cardiac Electrophysiological models

Modelling cardiac electrophysiology as well as trying to reproduce it as closely as possible to reality, elicit many interests such as: studying the mechanisms of arrhythmia initiation; improving our understanding of pathologies; model-based planning and guidance of treatment for better therapy outcome; etc. Several mathematical models have been developed to describe the genesis and propagation of the cardiac action potential wave, and these models can be categorised into: biophysical models; phenomenological models; and, eikonal models.

Biophysical models as [Tusscher, 2004], are complex models that consider different ionic concentrations at the cell level resulting in many parameters. These models require substantial computer resources (and often super-computers); however, they are very accurate in modelling the behaviour of heart cells. Phenomenological models are simplified models derived from biophysical models. Among them are [Fitzhugh, 1961; Aliev, 1996; Mitchell, 2003]. These models employ less parameters and therefore are computationally more attractive than the complex biophysical models. Therefore, they are especially useful for rapid computational modelling of wave propagation at organ level. Depending on the electrical potentials considered (extra-intra-cellular potential or transmembrane potential), there are two categories of phenomenological models:

bidomain and monodomain models. The bidomain model considers that the heart is formed by two sub-spaces: the intracellular and extracellular space, and that those compartments are both anisotropic. This model is very expensive in terms of computation times, but is very efficient and often considered as the reference method for cardiac electrical activity simulations; for more details see [Lines, 2003]. The monodomain model is a simplification of the bidomain model, under the assumption that extracellular and intracellular spaces are anisotropic, but with equal anisotropy ratio.

Eikonal models [Franzone, 1990; Keener, 1991; Jacquemet, 2011] are even more simplified models than phenomenological ones. A mathematical asymptotic expansion of the solution shows that the time of arrival of the depolarisation solves an Eikonal equation. As a result, these models are very fast in computation time, but it is complicated to simulate complex phenomena because of the lack of repolarisation times. However its associated results in terms of local activation times are comparable to more sophisticated models [Potse, 2006].

Considering the pros and cons of these different types of methods, in this thesis, we chose to use the Eikonal model from [Cedilnik, 2018] to simulate our cardiac activation maps. Further details regarding the Eikonal model are given in section 3.2.3.

1.3.2 Electrocardiographic Imaging

The electrical activity of the heart can be reconstructed in a non-invasive manner from electrocardiograms on the surface of the body, and from 3D images of the heart and from the torso. The modality allowing this panoramic three-dimensional visual reconstruction of the cardiac electrical activity is the ECGI. The first sketch has been proposed in 1977 [Barr, 1977]. Recently, its history and clinical applications were reviewed in [Pereira, 2020]. While important progress has been achieved, there are still scientific challenges due to the ill-posedness of the classical mathematical formulation. Furthermore, this emerging modality is still in early stages for the clinical practice due to its currently limited accuracy and resolution.

1.3.2.1 Forward problem

The forward problem of ECGI refers to the computation of the potentials on the torso surface knowing the source (i.e., the electrical activity of the heart). It could be decomposed into three parts: the input (which is the cardiac electrical activity); the output corresponding to the BSP; and, finally, the mathematical model to compute the propagation including assumptions on the electrical conductivity. The classical approach is based on a transfer matrix between epicardial potentials and the torso potentials. To compute this transfer matrix, different approaches were proposed such as: the Boundary

Element Method (BEM) [Shou, 2007; Stenroos, 2008; Wang, 2009a; Schuler, 2017] based on surfaces, or the Finite Element Method (FEM) [Wang, 2010; Zemzemi, 2015], whereas the 3D torso anatomy is approximated by small volume elements. Both methods propagate the epicardial action potentials to the body surface with chosen boundary conditions, *e.g.* null current across the body surface. Among many other methods that were proposed to compute this transfer matrix, the method of fundamental solution (MFS) [Wang, 2006; Chamorro-Servent, 2017] is a meshless approach.

In [Ramanathan, 2001], the authors demonstrated that ECGI reconstruction does not require the detailed inclusion of torso heterogeneities. In an *in vivo* porcine study [Bear, 2018b], the authors show that including inhomogeneities of the torso does not improve the inverse solution. In the context of our study, the direct model is applied to the simulation of BSP. These simulations are usually done on previously simulated cardiac activation maps and not on maps from EAM.

1.3.2.2 Inverse problem

The inverse problem is naturally based on the forward problem, but in this case the mathematical-physical relationship between BSP and electrical activity of the heart is reverted. Considering the input and output described above, here it is the output that is known and our goal is to find the input. Standard resolution is performed by estimating a transfer matrix from the following linear equation:

$$b = Ma \tag{1.1}$$

with b the Body Surface Potentials such as $b \in \mathbb{R}^m$, a the cardiac sources (either the transmembrane potential or the surface potential) such as $a \in \mathbb{R}^n$ and M the transfer matrix which depends on the model chosen and which belongs to $\mathbb{R}^{n \times m}$.

The cardiac source models used to solve the inverse ECGI problem are most often surface-based models (*e.g.* [Huiskamp, 1989; Greensite, 1998; Serinagaoglu, 2006; Ghosh, 2009]). Instead of estimating the potential at the surface of the heart, methods reconstructing the potential throughout the 3D-myocardium have been investigated by several groups [He, 2003; Wang, 2009b]. Other methods [Liu, 2006; Han, 2015] solve the inverse problem by estimating the equivalent current density instead of the TMP.

However, in most formulations, the inverse problem is ill-posed [Zemzemi, 2015; Pullan, 2004] which means that its solution is not unique, and that its resolution will be very sensitive to noise. Small errors in input will cause large errors in the output. Thus, in addition to the three components of the direct model, there is a fourth component which aims to optimise the reconstructions by reducing this sensitivity to noise. Several formu-

lations and regularisation methods have been proposed to stabilise this problem, see for instance the publications of the ECGI consortium³. Two very clear and thorough reviews on the subject of regularisation methods are [Karoui, 2018; Milanič, 2014]. Recently, in [Karoui, 2018], 15 different algorithms to solve the inverse problem were compared and the conclusion was that each of them had benefits which varied according to bi-/ , left-/ and right-ventricular pacing. Moreover, to solve the ECGI inverse problem, it is possible to include a multitude of physiological, physical and mathematical priors for the forward and inverse problem but also for describing the cardiac sources, e.g. activation map in [Van Oosterom, 2005], as well as to perform the ECG signal processing [Bear, 2018a]. An overview of most of these methods was published recently [Cluitmans, 2018].

However, ECGI could achieve more accurate and robust results by leveraging all the information present in the different imaging modalities acquired, as well as by introducing prior knowledge from previous cases.

1.3.3 Brief review of Deep Learning models and Related work

Being aware of the various advantages of Deep Learning (DL), many studies based on DL have been developed for solving the ECGI inverse problem. In particular, notable are the data-driven regularisation approaches such as: time-delayed Neural Network [Karoui, 2019]; Relevant Vector Regression [Giffard-Roisin, 2018]; and, Variational Autoencoder [Ghimire, 2018; Ghimire, 2019].

The method developed during this thesis in order to solve the whole inverse problem and reconstruct accurate 3D activation maps is based on DL and autoencoders. More specifically, we propose to reformulate the whole ECGI problem as a β -conditional variational autoencoder based on convolutional neural networks.

1.3.3.1 Introduction to Deep Learning models

Deep Learning (DL) is a subfield of Machine learning (ML), which itself is a branch of Artificial Intelligence (AI). The term "Artificial Intelligence" (AI) is referred to the science of making machines intelligent, and was first introduced in 1956 by Professor McCarthy. ML models are based on mathematical optimisation and statistics and use data to learn what they are designed for. Usually, the data are preprocessed and split into three subsets: training set; validation set; and, test set. The ML model is trained to learn from the training set, the validation test allows us to tune hyperparameters and avoid overfitting, and is evaluated using the test dataset. Deep learning is based on Neural Networks, which can be referred to human neurobiology. A neural network (NN) is a succession of

³<http://www.ecg-imaging.org>

layers that are linked together and can be optimised into specific types of patterns. Each layer has an associated weight. The weight parameterises the layer, thus influences the transformation that will be applied to the input data. The transformed output of this layer is then passed to the subsequent layer, where a different transformation is done according to its weight. In general, for the model to be as efficient as possible, these weights must be optimised. Thus, in the training phase of the model, a loss function is calculated as the error distance between the predictions and the expected values. Logically, the smaller the difference, the better the network is at solving its task. This loss is back-propagated through each trainable layer to generate the gradients. Then, back-propagated gradients are used to update the weights of each layer using the optimiser function.

Let us introduce some other key concept:

- **Overfitting and underfitting:** both prevent the learning algorithm from generalizing to new data. Overfitting occurs when the DL model has learned too specific patterns from the training data and its patterns are irrelevant to new, unseen data. Conversely, underfitting is an under-specialisation of the training data.
- **Supervised and unsupervised learning:** supervised learning algorithm is trained with labelled data. In contrast, unsupervised learning algorithm work on its own to discover interesting pattern as it is trained without labelled data.
- **Hyperparameters and parameters:** Hyperparameters are fixed by the user and remain unchanged throughout the learning process. Parameters are updated throughout the training process until they ideally reach the values that make the model as accurate as possible.
- **Epoch:** During the training phase, an epoch represents a cycle for which the model has visualised the complete training set.
- **Batch:** In practice, the hyperparameter we define is the batch size, and corresponds to the size of a portion of the whole dataset which will be seen during an epoch. Thus, for an epoch the batch is equivalent to the number of total training data divided by the batch size.
- **Optimiser:** It is a mathematical function that depends on the learnable parameters of the DL model and that is used in order to reduce the loss function. A wide variety of optimizers exist (e.g. Adam) and each has its own way of modifying the learnable parameters, such as the weights.

- Learning rate: The learning rate is a hyperparameter of the optimiser, and is used to regulate the frequency of updating the weights. Although it is a 'hyperparameter', its value is not necessarily fixed but can vary, depending on how the user defines it. Notably, this is one of the most important hyperparameters to be tuned for the model.

Different types of layers exist within an NN, each with its own specificity. Perhaps the most popular and well-known of the deep learning algorithms is the convolutional neural network (CNN).

1.3.3.2 Convolutional neural networks

Convolutional neural networks [Lecun, 1998] are NN in which most layers are convolution layers. These layers are so called because they perform a mathematical operation: convolution. Convolutional neural networks CNNs often combine different types of layers in succession: a convolutional layer followed by an activation (often *ReLU*) and a pooling layer. The convolution operation acts like a filter that slides over the data and extracts features (see an illustration of it in Figure 1.6a). The dimensions of the filter and the step size with which it scans the data are to be defined. The result of a convolution operation is a feature map characterising the input data of the layer. Each convolution layer applies several convolutions, so several feature maps are created. The activation layer consists in applying a non-linear function to the data it receives as input (i.e., the feature maps). This allows us to extract complex features. The choice of the non-linear function used depends on the data being processed and the task expected by the model. Finally, pooling is a sub-sampling method. Figure 1.6b shows a Max-pooling operation. Most often, the pooling layer is between the activation and convolution layers, allowing us to reduce the size of the feature map and therefore to reduce not only the computation time but also to reduce the overfitting risk. A typical CNN model architecture is represented in Figure 1.6c.

Convolutions allow spatial information to be propagated through all layers of the CNN. When data has underlying temporal and spatial information, it is often chosen to use convolutions and 3D pooling which together allow not only spatial information but also temporal information to be propagated throughout the network. However, it is not necessary for all convolution layers to be 3D. It can be a mixture of 1D (to capture temporal information), 2D (to capture spatial information) and 3D (for spatio-temporal information) convolution layers. Thus, thanks to the convolution operations, CNNs are able to capture spatial and temporal dependencies from data such as BSP and cardiac activation data.

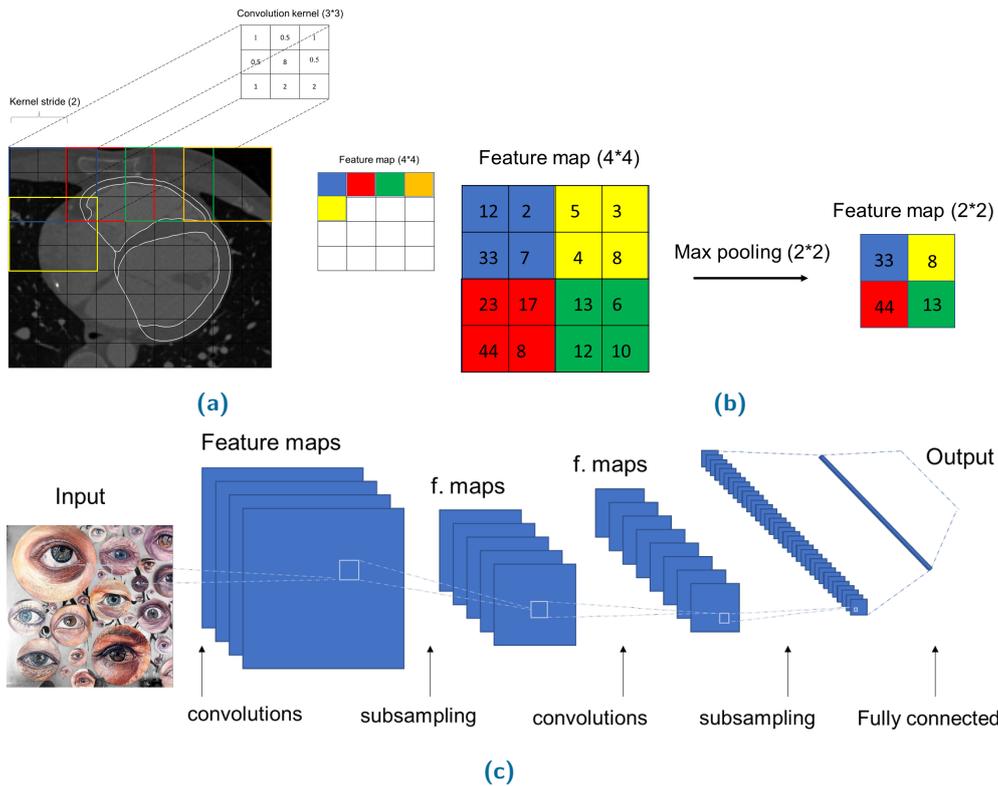


Fig. 1.6: (a) Diagram of a convolution with a kernel size 3x3 and a stride of 2, (b) Diagram of a pooling operation with a maxpool kernel of 2x2 and stride of 2, (c) shows a typical CNN model architecture.

1.3.3.3 Autoencoder

An autoencoder [Kingma, 2014b] is an unsupervised learning model whose aim is to reconstruct output data from a latent space to be as close as possible to the input data. In order to achieve this goal, the autoencoder tries to learn the best possible representation of the data. It consists of two connected networks: an encoder and a decoder (as shown in Figure 1.7). The encoder takes an input x and compresses it into a low-dimensional latent representation known as the bottleneck layer, which variables are denoted z . Then the decoder reconstructs from z the output data x' . The indicator of the performance of the model is the reconstruction loss, i.e. the difference between x and x' . The smaller the reconstruction loss, the better the model. This is possible when the latent space represents the most important feature variations of the original data.

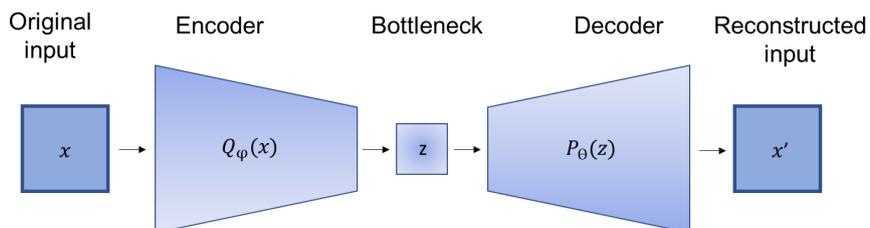


Fig. 1.7: Illustration of autoencoder model architecture.

1.3.3.4 Variational autoencoder

Bayes statistics methods are often used to compute posterior probability distributions in order to do statistical inference. Variational Inference (VI) is used to approximate these probability distributions, which are difficult to calculate by optimisation. As the autoencoder, the VAE [Kingma, 2014b] consists of two connected networks: an encoder which is known as the inference model of VAE and a decoder that is the generative model. The former takes an input x and compresses it into a low-dimensional latent representation, in which variables are denoted z , with a prior distribution $P(z)$. Thus, unlike the autoencoder, the VAE encoder describes a probability distribution for each latent attribute. The prior distribution $P(z)$ is often assumed to be a centred isotropic multivariate Gaussian $P(z) = N(z : 0, I)$. The decoder takes z as input and reconstructs the data from the generative distribution $P_{\Theta}(x|z)$. This likelihood distribution $P_{\Theta}(x|z)$ is learned in the decoder neural network with parameters Θ . The resulting generative process induces the distribution $P_{\Theta}(x) = E_{P(z)}P_{\Theta}(x|z)$. Due to the intractability of the true posterior distribution $P(z|x)$, the posterior is approximated by learning a simpler probability $Q_{\Phi}(z|x)$ in the encoder neural network with parameters Φ . A VAE is trained in order to maximise the probability of generating real data samples, by maximising the Evidence Lower Bound (ELBO). The ELBO is composed of two terms: the Kullback-Leibler (KL) divergence which is a measure between the variational probability $Q_{\Phi}(z|x)$ and the true posterior distribution $P(z|x)$; and, the log likelihood of $P_{\Theta}(x)$ under the probability distribution $P_{\Theta}(x|z)$. By minimising the loss function, the ELBO is maximised.

$$L_{\text{VAE}}(\Theta, \Phi) = -\log P_{\Theta}(x|z) + D_{KL}(Q_{\Phi}(z|x) \| P(z)) \quad (1.2)$$

with the Kullback-Leibler divergence:

$$D_{KL}(Q_{\Phi}(z|x) \| P(z)) = \sum Q_{\Phi}(z|x) \log \frac{Q_{\Phi}(z|x)}{P(z)} \quad (1.3)$$

Commonly, $Q_{\Phi}(z|x)$ is Gaussian with a diagonal covariance matrix: $z \sim Q_{\Phi}(z|x) = N(z : \mu, \sigma)$ where: $z = \mu + \sigma \odot \epsilon$, \odot denotes the Hadamard product and $\epsilon \sim N(0, I)$. The D_{KL} term can be computed analytically (see appendix B in [Kingma, 2014b]).

1.3.3.5 Deep probabilistic generative model applied to ECGI

The VAE model is a probabilistic generative model; however, as it stands, it does not allow us to solve the ECGI inverse problem. Indeed, as the inference model models z directly based on x , it allows us only to generate the data of the modality that we give it as input. For example, if x represents BSP data, the model will generate new BSPs (see section 2.3.2.1). What we expect for our model is to be able to reconstruct new

cardiac activation maps knowing BSP and the cardiac geometry. In other word, the cardiac electrical activity x is conditioned to its two corresponding data which will be noted as c . This model is a Conditional Variational AutoEncoder (CVAE) [Kingma, 2014c; Sohn, 2015]. The encoder of the CVAE has therefore the data x and the conditioning data c as input, which results in the variational probability: $Q_{\Phi}(z|x, c)$. c modulates also the prior distribution. Respectively, the decoder also has as input the conditioning data c . Thus, the generative distribution becomes: $P_{\Theta}(z|x, c)$. Figure 1.8 illustrates the architecture of a CVAE model.

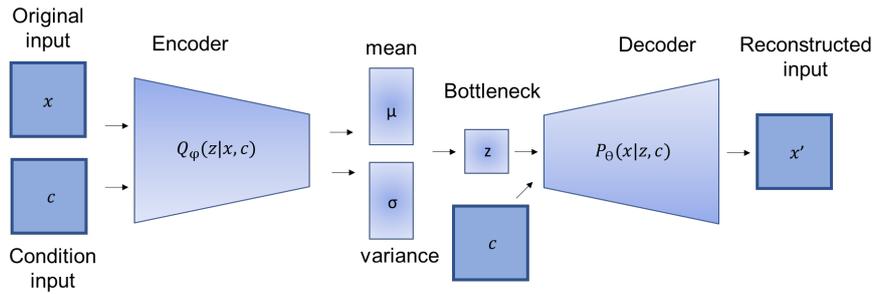


Fig. 1.8: Structure of a CVAE model.

In [Higgins, 2017], it is shown that introducing a penalty coefficient β on the KL divergence, the model obtains a better disentangling performance than conventional VAE. This allows it to discover disentangled latent factors. With this in mind, we have opted to add this factor to our CVAE. Specifically, in this work, we propose a β -CVAE model which loss function is therefore defined by a reconstruction term $P_{\Theta}(x|z, c)$ and a weighted KL divergence between variational and prior probabilities term. As this approach is based on VAE, the β -CVAE is also trained to minimise the loss function:

$$L_{\beta\text{-CVAE}}(\Theta, \Phi, \beta) = -\log P_{\Theta}(x|z, c) + \beta D_{KL}(Q_{\Phi}(z|x, c) \| P(z)) \quad (1.4)$$

with β the penalty coefficient, $Q_{\Phi}(z|x, c)$ is Gaussian with a diagonal covariance matrix: $z \sim Q_{\Phi}(z|x, c) = N(z : \mu, \sigma)$ where: $z = \mu + \sigma \odot \epsilon$, \odot denotes the Hadamard product and $\epsilon \sim N(0, I)$.

1.3.3.6 State-of-the-art of Deep Learning models for ECGI

Although very promising, to our knowledge, the first DL-based model for solving the inverse problem of ECGI was only proposed in 2010 by [Jiang, 2010]. Since then, different methods based on deep learning have been proposed, but nevertheless there are only few. In [Jiang, 2010] support vector regression (SVR) [Vapnik, 1999] is used to reconstruct TMP at the epicardial surface of the heart. The reconstructions produced by this method are more accurate than those produced by the Tikhonov regularisation. The same author in [Jiang, 2011] improves his previous SVR-based method by introducing the kernel principal component analysis (KPCA) to extract features from input data. In [Jiang,

2013] a clustering model using self-organising feature map [Kohonen, 2003] techniques is combined with the SVR model. This two-steps method performs better the reconstruction than the previous ones and also reduces the computation time. A kernel ridge regression model is proposed in [Zemzemi, 2012]. This model is based on Hilbert space and shows its ability to estimate cardiac activation maps on the surface of the atria from simulated BSP. However, the model is very dependent on the richness of the training dataset. Our group [Giffard-Roisin, 2018] previously proposed a method using a reference torso anatomy onto which the BSPs were projected before using an relevance vector regression method (to solve the inverse problem), and the reconstructions obtained were the cardiac activation maps. Others [Ghimire, 2018] developed a model to reconstruct the TMP based on a VAE with long short-term memory (LSTM) networks. By comparing the reconstructions obtained with a method constrained by conventional physiological models, it was shown that the proposed model introducing Bayes inference performed better than the latter for both simulated and clinical data. Furthermore, methods based on Time-delay neural network have been investigated in [Malik, 2018; Karoui, 2019]. For instance, the study in [Malik, 2018] focused on the prediction of a single heart surface potential from a BSP, and the results obtained proved to be acceptable with clinical data. This is undoubtedly what motivated [Karoui, 2019] to improve the method and to further generalise it in order to perform reconstruction at the whole epicardial surface scale. Moreover, a recent DL approach based on the ISTA (Iterative Soft Thresholding) algorithm was proposed by Wang and colleagues [Wang, 2019], who showed that their reconstructed potentials on the cardiac surface are of superior quality than the one reconstructed with three different regularisation methods, including Tikhonov-based method. Another advantage of their approach is the speed of convergence. However, the cardiac potentials seem to have been estimated on the same geometry as the one used for training the model. In addition, this method, like those presented in this paragraph, solves the problem for each time step independently.

1.4 Conclusion

In this chapter, we provide a clinical and scientific background to cardiac electrophysiology. First, we introduced the cardiac anatomy and its electrical functioning and described the clinical modalities of its study. After summarising the cardiac electrophysiological models, Electrocardiography was highlighted, its assets and challenges were presented. Finally, we provided a brief review of Deep Learning models and compiled a state-of-the-art of DL based models in the field of Electrocardiography. In this thesis, we focus specifically on the development of a new DL-based approach to tackle the inverse problem of Electrocardiography.

Evaluation in 2D on Synthetic Data

Contents

2.1	Introduction	26
2.2	Context	27
2.2.1	Electrocardiographic Imaging	27
2.2.2	Deep Learning	27
2.3	Methods	30
2.3.1	ECGI Forward Problem: Data Simulation	30
2.3.2	Models Structure and Training	30
2.4	Results	32
2.4.1	Evaluation of Body Surface Potentials VAE	32
2.4.2	Evaluation of Activation Maps VAE	33
2.4.3	Evaluation of β -CVAE	33
2.5	Discussion & Conclusion	34

This chapter is based on the following paper:

Deep Learning Formulation of ECGI for Data-Driven Integration of Spatiotemporal Correlations and Imaging Information. *Bacoyannis T., Krebs J., Cedilnik N., Cochet H., Sermesant M.* In Functional Imaging and Modelling of the Heart (FIMH) 2019.

Abstract: The challenge of non-invasive Electrocardiographic Imaging (ECGI) is to re-create the electrical activity of the heart using body surface potentials. Specifically, there are numerical difficulties due to the ill-posed nature of the problem. We propose a novel method based on Conditional Variational Autoencoders using Deep generative Neural Networks to overcome this challenge. By conditioning the electrical activity on heart shape and electrical potentials, our model is able to generate activation maps with good accuracy on simulated data (mean square error, MSE = 0.095). This method differs from other formulations because it naturally takes into account spatio-temporal correlations as well as the imaging substrate through convolutions and conditioning. We believe these features can help improving ECGI results.

Keywords: ECGI·Deep learning·Simulation·Generative Model

2.1 Introduction

Electrocardiographic Imaging (ECGI) has been an active research area for decades. Important progress was achieved but there are still challenges in robustness and accuracy due to the ill-posedness of the classical formulation.

In the last few years, deep learning (DL) based methods have been used to solve inverse problems, e.g. in medical image reconstruction [Hammernik, 2018]. Autoencoders are popular for these problems, as they are specifically designed to reconstruct high dimensional data in an unsupervised fashion. Autoencoders learn an identity function in order to reconstruct the input image after having first encoded it in a latent representation and then decoded it to the original input. Such methods have been recently introduced into the ECGI problem to regularize the temporal information while processing it [Ghimire, 2018].

In this manuscript, we propose to reformulate the whole ECGI problem as a conditional variational autoencoder based on convolutional neural networks. This has four main advantages:

- Spatio-temporal correlations: the convolutional model learns interactions in space and time between signals, while most ECGI methods solve each time step independently.
- Imaging substrate: the correlation between the substrate from imaging and the signals is also learned, therefore we can seamlessly integrate any 3D image information in ECGI, while this is still difficult in the classical formulation.
- Data-driven regularisation: using a generative model from a low dimensional space should ensure smooth variations between similar cases; this could alleviate the ill-posedness problem.
- Fast computations: once trained, DL methods are very fast to evaluate.

In order to achieve this, we use Cartesian coordinates for all the data (space and time) to leverage the power of convolutional neural networks.

2.2 Context

2.2.1 Electrocardiographic Imaging

ECGI is a non invasive modality which aims to better understand the electrical activity of the heart, both quantitatively and qualitatively. ECGI allows the visualisation of the electrical potential distribution of the electrical wave on the heart surface from body surface potentials (BSP). It requires medical imaging to obtain geometrical information and methods to solve the inverse problem.

2.2.1.1 Forward Problem

This refers to the estimation of the ECG data from cardiac data. The two classical numerical approaches for this are the Boundary Element Method (BEM), based on surfaces, and the Finite Element Method (FEM), where the 3D torso model is approximated by small volume elements. Both propagate the epicardial action potentials to the body surface with chosen boundary conditions, *e.g.*, null current across the body surface. There are also methods using a dipole formulation, assuming that the torso domain is homogenous and infinite [Chavez, 2017]. In [Ramanathan, 2001], the authors demonstrated that noninvasive ECGI reconstruction does not require first order approximations for torso heterogeneities.

2.2.1.2 Inverse Problem

This allows to reconstruct cardiac electrical activity using BSP. The classical approach is based on a transfer matrix between epicardial potentials and the torso potentials. To compute this transfer matrix different approaches can be used such as BEM [Stenroos, 2008], FEM [Zemzemi, 2015] or the method on fundamental solution (MFS) [Chamorro-Servent, 2017]. However, inverting it is ill-posed [Zemzemi, 2015]. Therefore, different formulations and regularization methods were proposed, see for instance the publications of the ECGI consortium¹.

2.2.2 Deep Learning

2.2.2.1 Variational Autoencoders

Variational Autoencoders (VAE) are powerful probabilistic generative models [Kingma, 2014b]. VAEs consist of two connected networks: an encoder and a decoder. The former

¹<http://www.ecg-imaging.org>

takes an input x and compresses it into a low-dimensional latent representation, which variables are denoted z , with a distribution $P(z)$, often assumed to be a centred isotropic multivariate Gaussian $P(z) = N(z : 0, I)$. The latter takes z as input and reconstructs the data from the generative distribution $P_{\Theta}(x|z)$. This likelihood distribution $P_{\Theta}(x|z)$ is learned in the decoder neural network with parameters Θ . The resulting generative process induces the distribution $P_{\Theta}(x) = E_{P(z)}P_{\Theta}(x|z)$. The term Variational in Variational Autoencoder refers to variational inference or variational Bayes. Due to the intractability of the true posterior distribution $P(z|x)$, the posterior is approximated by learning the probability $Q_{\Phi}(z|x)$ in the encoder neural network with parameters Φ .

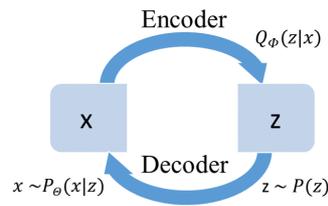


Fig. 2.1: VAE model [Kingma, 2014b]

A VAE is trained in order to minimize the Kullback-Leibler (KL) divergence between the variational probability $Q_{\Phi}(z|x)$ and the true posterior distribution $P(z|x)$. As this is intractable, it can be reformulated as the evidence lower bound (ELBO) of the log marginalized data likelihood $P_{\Theta}(x)$ [Kingma, 2014b]:

$$\log P_{\Theta}(x|z) - D_{KL}(Q_{\Phi}(z|x) \parallel P(z)) \leq \log P_{\Theta}(x)$$

The VAE loss function is then defined by a reconstruction term and the KL divergence between variational and prior probability:

$$L_{\text{VAE}}(\Theta, \Phi) = -\log P_{\Theta}(x|z) + D_{KL}(Q_{\Phi}(z|x) \parallel P(z))$$

The Kullback-Leibler divergence is:

$$D_{KL}(Q_{\Phi}(z|x) \parallel P(z)) = \sum Q_{\Phi}(z|x) \log \frac{Q_{\Phi}(z|x)}{P(z)}$$

By minimizing the loss function, the lower bound of the probability of generating real data samples is maximised. To compute the gradient of the variational lower bound, we use the reparameterization trick with respect to VAE [Kingma, 2014b]. Commonly, $Q_{\Phi}(z|x)$ is Gaussian with a diagonal covariance matrix: $z \sim Q_{\Phi}(z|x) = N(z : \mu, \sigma)$ where $z = \mu + \sigma \odot \epsilon$ and $\epsilon \sim N(0, I)$.

2.2.2.2 Conditional Variational Autoencoders

Conditional VAEs (CVAEs) [Kingma, 2014c] are an extension of VAEs [Kingma, 2014b] where the latent variables and the data are both conditioned on additional random variables c . The encoder of the CVAE is not only conditioned on the data x but also on the conditioning data c which results in the variational probability: $Q_{\Phi}(z|x, c)$. Respectively, the decoder is also conditioned on the conditioning data c . Thus, the generative distribution becomes: $P_{\Theta}(z|x, c)$. The CVAE is trained to maximize the conditional log-likelihood where the ELBO is:

$$L_{\text{CVAE}}(\Theta, \Phi) = -\log P_{\Theta}(x|z, c) + D_{KL}(Q_{\Phi}(z|x, c) \parallel P(z)) \leq \log P_{\Theta}(x|c)$$

And the KL divergence is:

$$D_{KL}(Q_{\Phi}(z|x, c) \parallel P(z)) = \sum Q_{\Phi}(z|x, c) \log \frac{Q_{\Phi}(z|x, c)}{P(z)}$$

2.2.2.3 β -Variational Autoencoders

β -VAEs are an evolution of VAEs which intends to discover disentangled latent factors. In [Higgins, 2017], it is shown that this model achieves similar disentanglement performance compared with VAEs, both quantitatively and qualitatively. β -VAEs and VAEs have the same goals: maximize the probability of generating real data and minimize the distance between the real and estimated posterior distributions (smaller than a constraint ε). In this approach, the prior is an isotropic Gaussian $P(z) = N(0, I)$.

The β -VAE lower bound is defined as:

$$L_{\beta\text{-VAE}}(\Theta, \Phi, \beta) = -\log P_{\Theta}(x|z) + \beta D_{KL}(Q_{\Phi}(z|x) \parallel P(z))$$

with β a penalty coefficient. If $\beta = 1$, β -VAEs correspond to the original VAEs [Kingma, 2014b]. If $\beta > 1$, β -VAEs apply a stronger constraint on the latent bottleneck and so limit the capacity of z . It allows to learn the most efficient representation of the data. However, if β is too big, β -VAE learns an entangled latent representation because of its excessive capacity in the latent z bottleneck. Same remarks if β is too small, it will have too little capacity. To sum up, $\beta > 1$ is capital to achieve good disentanglement.

2.3 Methods

2.3.1 ECGI Forward Problem: Data Simulation

We simulated cardiac activation maps and BSP data using the Eikonal Model directly on a Cartesian grid from image segmentation [Cedilnik, 2018]. The Eikonal model is a fast generic model of wave front propagation able to reproduce any periodic activation map. Its inputs are the myocardial wall mask, a local conduction velocity v for each voxel x of the given wall mask and the pacing zone.

$$v(x) \|\nabla T(x)\| = 1$$

with $T(x)$ the local activation time in x . It is solved using The Fast Marching Method (FMM) [Sermesant, 2005].

BSP were generated using the dipole formulation [Giffard-Roisin, 2017] associating each activation time from the Eikonal model with an action potential signal from the Mitchell Schaeffer model. 100 torso electrodes were positioned on a 10×10 Cartesian grid in front of the heart.

In total, 120 activation maps (considering single pacing, sampled randomly from a discrete uniform distribution) and the corresponding BSP were simulated using a patient image segmentation from CT images. In order to accelerate computation and ease memory requirements, we currently present 2D results, where we separated the 3D simulated data in 2D slices.

2.3.2 Models Structure and Training

The different models described below were trained on 80% of the data, tested on the remaining 20% and have been implemented using Keras. ².

2.3.2.1 VAE for BSP

It follows the architecture presented in Figure 2.2, with convolutional layers in order to extract spatiotemporal correlations.

The encoder consists of two convolutional layers followed by one dense layer. The bottleneck layers (μ, σ, z) are fully-connected. The decoder consists of a fully-connected and three transposed convolutional layers. For all convolutional layers except the output

²<https://keras.io/>

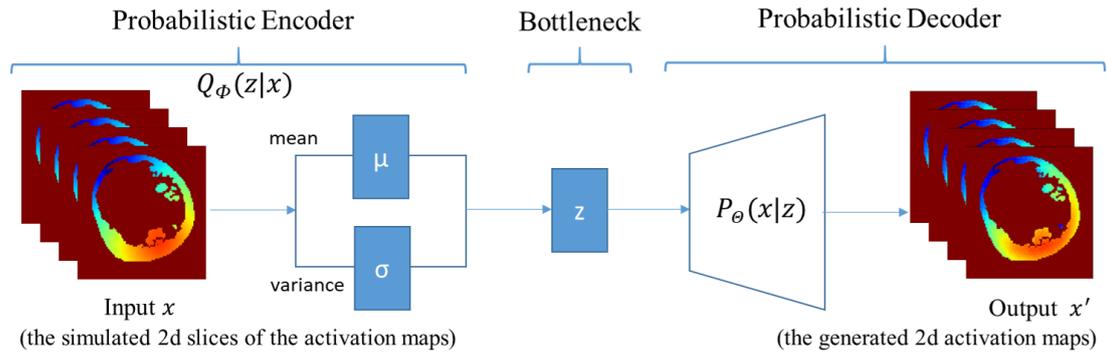


Fig. 2.2: Variational Autoencoder architecture

one: strides are set to 2, 16 filters are applied, the kernel size is 3 and the activation functions are *ReLU*. In the last layer, one filter is applied with a kernel size of 3 and a sigmoid activation function. The latent code size is set to 6. Mean squared error (MSE) is used to calculate the reconstruction loss. The optimization is performed with the NADAM solver with a batch size of 32.

2.3.2.2 VAE for Activation Maps

The encoder architecture is the same as the previous one. However, after tuning, the network parameters and activations differ: the convolution kernel size is 3 (even in the last layer) and all activations are *tanh*. The latent code size is set to 6. MSE is used to calculate the reconstruction loss. Optimization is performed with the RMSPROP solver with a batch size of 25.

2.3.2.3 β -CVAE for Activation Maps from BSP and Images

Our conditional autoencoder aims to generate heart activation maps using two conditions: the simulated ECG signals and the shape of the patient's heart. This is a novel formulation of ECGI using DL in order to learn the influence of cardiac shape/structure.

The architecture of our conditioned generative model (encoder) and our conditioned variational approximation (decoder) is described in Figure 2.3. Empirically, we found that the best value for Lagrangian multiplier hyperparameter β in the loss function was 5. The ADAM optimiser was used with 10^{-4} as learning rate and a batch size of 1. Latent space is set to 16.

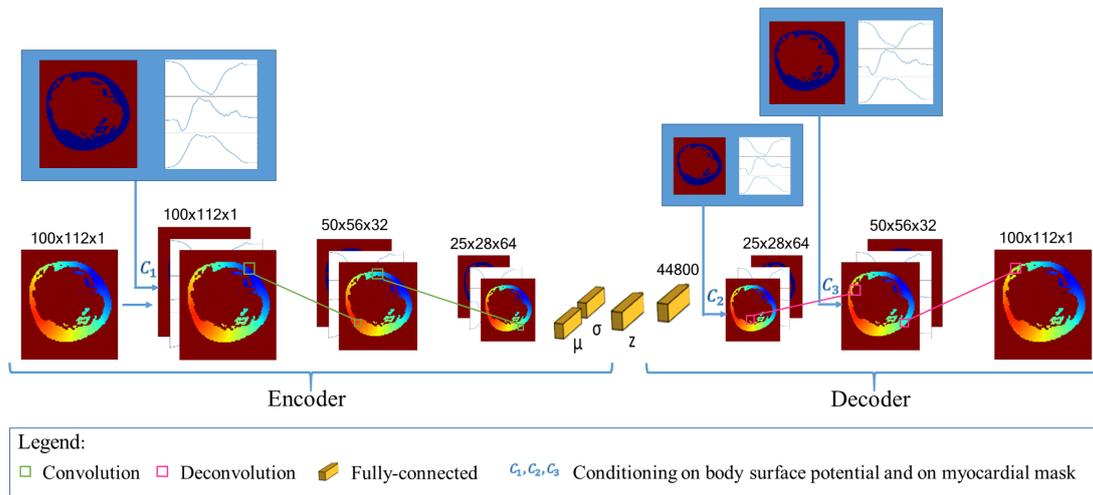


Fig. 2.3: CVAE architecture for ECGI with imaging data (C_1 : The first convolution layer in the encoder was conditioned by concatenating the input data with the BSP mapping signals and the myocardial mask. C_2, C_3 : The deconvolution layers in the decoder were conditioned by concatenating each layer's output with sub-sampled versions of the BSP mapping signals and of the mask).

2.4 Results

2.4.1 Evaluation of Body Surface Potentials VAE

Figure 2.4 shows two examples of input BSP mapping signals and their corresponding generated signals by our model.

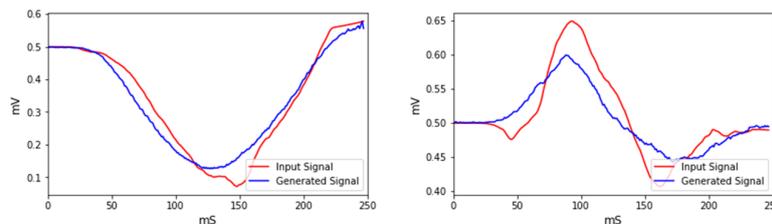


Fig. 2.4: Evaluation of input (red) and output (blue) signals from VAE for 2 different electrodes.

The performance of our model was evaluated measuring 5 metrics between the input BSP signals and their corresponding generated signals. We found a correlation of 0.95, a mean difference of 3.345%, an abscisse area difference of -2.152% and a maximum amplitude difference of -3.468%; 90.854% of the time, the sign of first peak was the same in the input and in the output.

2.4.2 Evaluation of Activation Maps VAE

Using our VAE model developed to generate the electrical activation of the heart (Figure 2.5), we obtained a MSE of 0.018 for reconstruction accuracy.

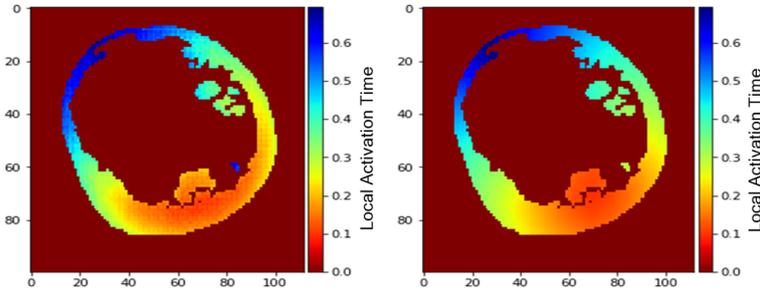


Fig. 2.5: Predicted (left) and simulated (right) activation maps with the Activation Maps VAE model.

2.4.3 Evaluation of β -CVAE

β -CVAE is a probabilistic generative method, so we can generate several probable solutions for a given input. We generated ten activation maps per prediction in order to evaluate the accuracy and stability of our method. The MSE metric was 0.095 over all the tests. Figure 2.6 shows an example of (a) an input activation map, (b) its corresponding generated mean activation map by our β -CVAE proposed method, (c) standard deviation map for 10 predictions, and (d) the error map.

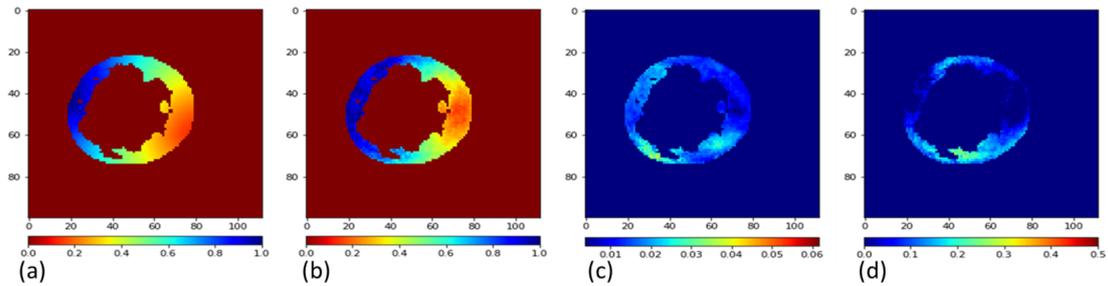


Fig. 2.6: (a) Simulated and (b) predicted mean activation maps for proposed deep learning based ECGI, (c) Standard deviation map calculated over 10 predictions, (d) error map, difference between predicted and simulated activation maps.

In Figure 2.6 (c) and (d), small values imply that the reconstruction performs well (small error and small standard deviation), while large values mean that the reconstruction is suffering. We can observe that the areas with the highest standard deviation are close to areas with the highest error, therefore the probabilistic aspect of our method can help in quantifying the uncertainty in the predictions.

2.5 Discussion & Conclusion

As a direct application of the Forward Problem of ECGI, electrical activation and potentials of the heart were simulated using a patient CT-scan image and the Eikonal Model. The first step of our work was to understand and choose the best hyperparameters of the Variational Autoencoders to generate potential or activation maps. Finally, we proposed a novel method based on Conditional β Variational Autoencoder able to solve ECGI inverse problem in 2D. This generative probabilistic model learns geometrical and spatio-temporal information and enables to generate the corresponding activation map of the specific heart. We showed that these generated electrical activities were very similar to the simulated ones. The presented method will now be generalised to 3D and evaluated on clinical data. Theoretically there is no impediment to extend our 2D model to 3D, as all the convolution operators have a 3D version. However, it will require to simulate more data to train the model, as we will have more hyperparameters to optimise. We will also explore transfer learning to apply it on clinical data [Giffard-Roisin, 2018].

Generalisation to 3D and Unseen Cardiac Geometries

Contents

3.1	Introduction	36
3.2	Methods	39
3.2.1	Deep Learning Approach	39
3.2.2	Supervised Learning Loss Function	42
3.2.3	Synthetic Data Generation: ECGI Forward Problem	43
3.2.4	Computational Details	45
3.3	Results	46
3.4	Discussion	47
3.5	Conclusion	48

This chapter is based on:

Deep learning formulation of electrocardiographic imaging integrating image and signal information with data-driven regularization. *Bacoyannis T., Ly B., Cedilnik N., Cochet H., Sermesant M.* In: EP Europace (2021), Volume 23, Issue *Supplement*₁, March 2021, Pages i55–i62.

Abstract:

Aims Electrocardiographic Imaging (ECGI) is a promising tool to map the electrical activity of the heart non-invasively using body surface potentials (BSP). However, it is still challenging due to the mathematically ill-posed nature of the inverse problem to solve. Novel approaches leveraging progress in artificial intelligence could alleviate these difficulties.

Methods We propose a Deep Learning (DL) formulation of ECGI in order to learn the statistical relation between BSP and cardiac activation. The presented method is based on Conditional Variational Autoencoders (CVAE) using deep generative neural networks. To quantify the accuracy of this method, we simulated activation maps and BSP data on six cardiac anatomies.

Results We evaluated our model by training it on five different cardiac anatomies (5 000 activation maps) and by testing it on a new patient anatomy over 200 activation maps. Due to the probabilistic property of our method, we predicted 10 distinct activation maps for each BSP data. The proposed method is able to generate volumetric activation maps with a good accuracy on the simulated data: the mean absolute error is 9.40 ms with 2.16 ms standard deviation on this testing set.

Conclusion The proposed formulation of ECGI enables to naturally include imaging information in the estimation of cardiac electrical activity from body surface potential. It naturally takes into account all the spatio-temporal correlations present in the data. We believe these features can help improve ECGI results.

Keywords: Electrocardiographic Imaging; Inverse Problem; Deep learning; Computational Modelling; Generative Model

What's New

- The proposed deep learning formulation of ECGI leverages the available cardiac imaging to better reconstruct the spatio-temporal electrical activity of the heart.
- This conditional generative model is able to learn spatio-temporal correlations as well as the influence of cardiac shape/structure through convolutions.
- As a deep learning model, its objective is to generalise the learned aptitude in order to predict accurate new activation for unseen data.
- Using a generative model from a low dimensional space should ensure smooth variations between similar cases; this could alleviate the ill-posedness issue and enables uncertainty quantification.
- The new ECGI formulation inherits the benefits of deep learning methods: computing power and memory capacity. Once trained, our model is fast to evaluate.

3.1 Introduction

Cardiac arrhythmias often require invasive catheter measurements in order to precisely diagnose the pathology and plan the therapy. This information on the local electrical activity in the myocardium can theoretically be obtained non-invasively through electrocardiographic imaging (ECGI). ECGI allows a panoramic three-dimensional visual

reconstruction of the electrical activity of the heart from body surface potentials (BSP), see Figure 3.1. This technique requires anatomical information of the heart and torso from 3D imaging, BSP measured with electrodes and their localisation on the torso surface, and mathematical methods to solve the inverse problem. ECGI has been an active research area for decades. The first sketch has been proposed in 1977 [Barr, 1977] and its history and clinical applications were recently reviewed [Pereira, 2020]. Important progress was achieved but there are still scientific challenges due to the ill-posedness of the classical mathematical formulation. Therefore, this emerging modality is still at its infancy in clinical practice due to its currently limited accuracy and resolution.

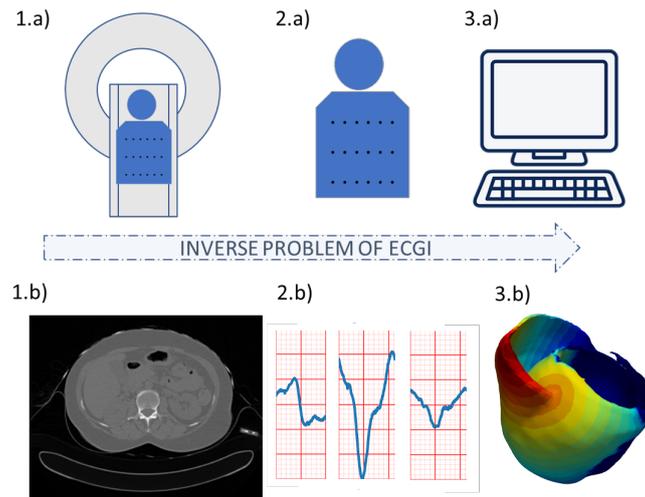


Fig. 3.1: ECGI process. Using 1.a) cardiac imaging and 2.a) signal processing, we obtain information on 1.b) the cardiac geometry and the position of the electrodes used to record the ECG signals, and 2.b) the surface ECG signals. Both are then used by 3.a) numerical methods to reconstruct the 3.b) cardiac electrical activity.

The classical approach is based on a transfer matrix between epicardial potentials and the torso potentials. The forward problem of ECGI refers to the computation of the electrical potential on the torso surface from the cardiac electrical activity. To compute this transfer matrix different approaches were proposed such as: the Boundary Element Method (BEM) [Wang, 2009a] based on surfaces, or the Finite Element Method (FEM) [Zemzemi, 2015] whereas the 3D torso anatomy is approximated by small volume elements. Both methods propagate the epicardial action potentials to the body surface with chosen boundary conditions, *e.g.* null current across the body surface. Another proposed approach is the method of fundamental solution (MFS) [Chamorro-Servent, 2017; Ramanathan, 2001], which is a meshless approach. In [Ramanathan, 2001], the authors demonstrated that ECGI reconstruction does not require the detailed inclusion of torso heterogeneities.

However, in most of the formulations, the inverse problem is ill-posed [Zemzemi, 2015] which means that its solution is not unique, and that its resolution will be very sensitive

to noise. Small errors in input will cause large errors in the output. Therefore, a variety of formulations and regularisation methods were proposed to stabilise this problem, see for instance the publications of the ECGI consortium¹. Recently, 15 different algorithms to solve the inverse problem were compared and the conclusion was that each of them had benefits which varied according to bi-/ , left-/ and right-ventricular pacing [Karoui, 2018]. Moreover, to solve the ECGI inverse problem, it is possible to include a multitude of physiological, physical and mathematical priors for the forward and inverse problem but also for describing the cardiac sources [Van Oosterom, 2005] as well as to perform the ECG signal processing [Bear, 2018a]. An overview of most of these methods was published recently [Cluitmans, 2018].

However, ECGI could achieve more accurate and robust results by leveraging all the information present in the different imaging modalities acquired, as well as by introducing prior knowledge from previous cases. This is what we propose in this manuscript, through the use of deep learning and generative models.

Deep learning (DL) has generated a huge interest in healthcare applications and has become the standard method for solving a large variety of computer vision problems. Moreover, these last years DL-based methods have proven their ability to solve inverse problems, e.g. in medical image reconstruction [Hammernik, 2018]. Recently, [Arridge, 2019] drew up a state-of-the-art of DL approaches for solving inverse problems, and has showed the benefits of such data-driven methods. [Ghimire, 2018] introduced autoencoders into ECGI for the regularisation of temporal information.

Our new numerical method is based on DL and autoencoders in order to solve the whole inverse problem and reconstruct accurate 3D activation maps. A preliminary version of this method in 2D and on a single geometry was presented in [Bacoyannis, 2019].

In this manuscript, we propose to reformulate the whole ECGI problem as a β -conditional variational autoencoder based on convolutional neural networks.

Using a DL model as autoencoder trained on a large database, i.e. activations with different pacing locations, allows to get more robust methods and reduces the prediction time. The developed method was used to predict activation maps for different cardiac geometries with their corresponding simulated BSP. The model was trained and tested on a simulated database using a personalised cardiac simulation pipeline [Cedilnik, 2019].

¹<http://www.ecg-imaging.org>

Compared with the classical mathematical formulation of ECGI, this approach has five main advantages:

- Spatio-temporal correlations: the convolutional model learns interactions in space and time between signals, while most ECGI methods solve each time step independently or use a temporal prior without spatial correlations.
- Imaging substrate: the correlation between the substrate from imaging (i.e. myocardial scar) and the signals is also learned, therefore we can seamlessly integrate any 3D image information in ECGI, while this is still difficult in the classical formulation.
- Data-driven regularisation: using a generative model from a low dimensional space should ensure smooth variations between similar cases; this could alleviate the ill-posedness issue.
- Volumetric predictions: this 3D convolutional approach generates volumetric activation maps within the whole myocardium, not limited to the epicardium and the endocardium.
- Fast computations: once trained, DL methods are very fast to evaluate.

In order to achieve this, we use Cartesian grids for all the data (space and time) to leverage the power of convolutional neural networks.

In the following section, we will first present our DL model, then the method used to simulate activation maps from CT images and we will describe the computation details of our model.

3.2 Methods

3.2.1 Deep Learning Approach

We base our reformulation of ECGI on a probabilistic generative model, namely a Conditional Variational AutoEncoder (CVAE) [Kingma, 2014c; Sohn, 2015]. A CVAE is an extension of a Variational AutoEncoder (VAE) [Kingma, 2014b] which is itself a powerful probabilistic generative version of autoencoders [Kingma, 2014b].

An autoencoder is an unsupervised learning technique which compresses the relevant information of the input data into a low-dimensional latent code in order to reconstruct output data with the least possible amount of difference with the input. This

dimensionality reduction allows to capture the more important features of the input data.

In a Variational AutoEncoder (VAE), the term Variational refers to variational inference or variational Bayes used to approximate complex distributions. As the autoencoder, the VAE consists of two connected networks: an encoder and a decoder. The former takes an input x and compresses it into a low-dimensional latent representation, which variables are denoted z , with a prior distribution $P(z)$, often assumed to be a centred isotropic multivariate Gaussian $P(z) = N(z : 0, I)$. The latter takes z as input and reconstructs the data from the generative distribution $P_{\Theta}(x|z)$. This likelihood distribution $P_{\Theta}(x|z)$ is learned in the decoder neural network with parameters Θ . The resulting generative process induces the distribution $P_{\Theta}(x) = E_{P(z)}P_{\Theta}(x|z)$. Due to the intractability of the true posterior distribution $P(z|x)$, the posterior is approximated by learning a simpler probability $Q_{\Phi}(z|x)$ in the encoder neural network with parameters Φ . A VAE is trained in order to maximise the probability of generating real data samples, by maximising the Evidence Lower Bound (ELBO). The ELBO is composed by two terms: the Kullback-Leibler divergence (as detailed in Equation 1.3), which is a measure between the variational probability $Q_{\Phi}(z|x)$ and the true posterior distribution $P(z|x)$; and the log likelihood of $P_{\Theta}(x)$ under the probability distribution $P_{\Theta}(x|z)$. By minimising the loss function, the ELBO is maximised.

In a Conditional Variational AutoEncoder (CVAE), the input data is conditioned on an additional random variable c that modulates the prior on the latent vector. The encoder of the CVAE has therefore the data x but also the conditioning data c as input, which results in the variational probability: $Q_{\Phi}(z|x, c)$. Respectively, the decoder also has as input the conditioning data c . Thus, the generative distribution becomes: $P_{\Theta}(z|x, c)$. Based on VAE, the CVAE is also trained to minimise the loss function:

$$L_{\text{CVAE}}(\Theta, \Phi) = -\log P_{\Theta}(x|z, c) + D_{KL}(Q_{\Phi}(z|x, c) \parallel P(z)) \quad (3.1)$$

Commonly, $Q_{\Phi}(z|x, c)$ is Gaussian with a diagonal covariance matrix: $z \sim Q_{\Phi}(z|x, c) = N(z : \mu, \sigma)$ where: $z = \mu + \sigma \odot \epsilon$, \odot denotes the Hadamard product and $\epsilon \sim N(0, I)$.

A β -VAE intends to discover disentangled latent factors by adding β , a penalty coefficient, on the KL divergence. In [Higgins, 2017], it is shown that with $\beta > 1$ the model achieves better disentanglement performance than classical VAEs.

We use a β -CVAE which loss function is therefore defined by a reconstruction term $P_{\Theta}(x|z, c)$ and a weighted KL divergence between variational and prior probabilities term:

$$L_{\beta\text{-CVAE}}(\Theta, \Phi, \beta) = -\log P_{\Theta}(x|z, c) + \beta D_{KL}(Q_{\Phi}(z|x, c) \parallel P(z)) \quad (3.2)$$

Artificial Neural Network Architecture

The proposed method aims to generate cardiac activation maps using two constraints: the 3D imaging data of the patient heart and the BSP signals. Therefore we use a multimodal input β -CVAE, which consists of three elements: the encoder, the latent vector and the decoder. The architecture of our deep learning model is described in Figure 3.2.

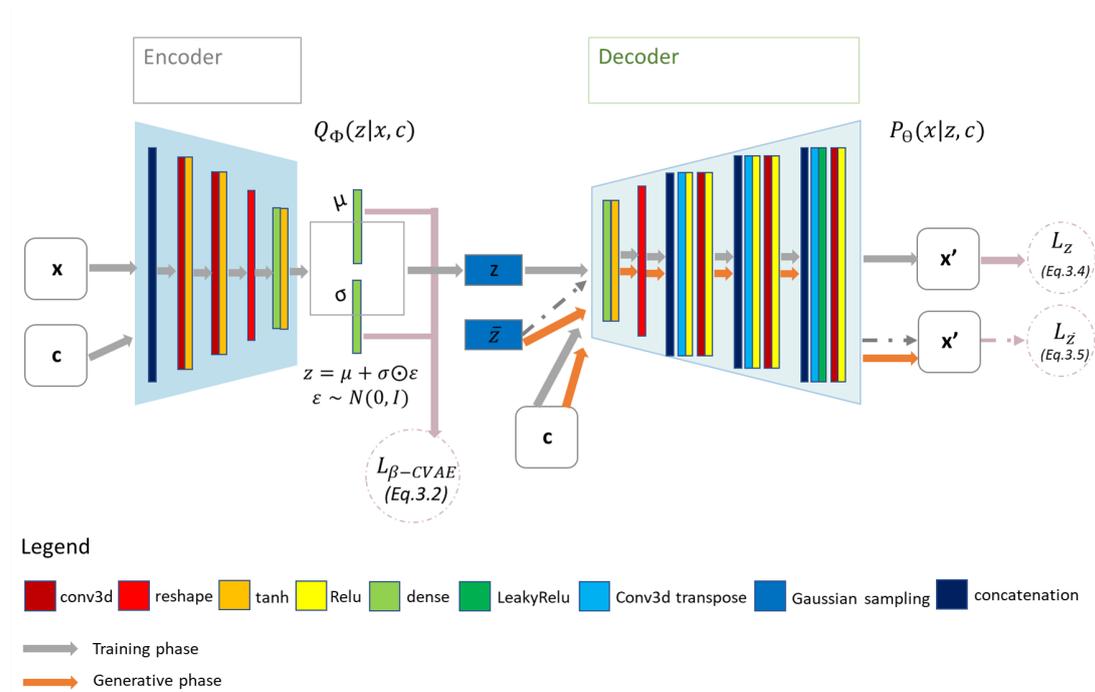


Fig. 3.2: β -CVAE model architecture. Training: input data is activation map concatenated with BSP mapping signals and with myocardial mask and encoded in z . Activation maps are decoded from z with conditioning by sub-sampled versions of the BSP mapping signals and of the cardiac mask. Prediction: decoder is used to generate myocardial electrical activation maps from a new cardiac segmentation and its corresponding BSP mapping signals.

First, the encoder takes as input data the normalised activation maps x and as conditioning data c the BSP signals and the cardiac image. As these data come from different modalities, BSP and imaging data do not have the same dimension. To obtain the same shape for the signals, we applied two successive 2D convolutional layers. Then these three input are concatenated to be used as a single input to the encoder model. This enables the network to directly learn spatiotemporal correlations within these input. The encoder consists of two successive convolutional layers (see subsection *Computation Details* for the implementation details) following the grey arrows, which compress the input.

The 3D output of these convolutional layers is flattened to 1D before being fed to a dense layer to further compress the output vector into 25 dimensions. We then generate from it

the multivariate Gaussian means and standard deviations μ and σ . In order to sample the latent vector z we used the Gaussian reparameterisation method.

The decoder network consists of a main branch which takes the vector z as input, and the conditional branches, where the conditional input c is integrated. The main branch applies a dense layer to the input z , before reshaping the 1D vector to a 3D matrix. Then, we use the combination of transpose and normal convolutional layers at each level to increase the output dimension back to its original size. At each level, we use two convolutional layers (strided followed by non-strided) on the condition inputs c to generate the condition input in the same shape as the output from the main branch. Then, we concatenate the condition input with the main branch output. At the end of the decoder, we multiply the output with the myocardial image binary mask M to only provide predictions within the myocardial wall.

3.2.2 Supervised Learning Loss Function

During the training of the CVAE, we use a modified version of the Root Mean Squared Error (RMSE), Eq. (3.3), to calculate the reconstruction loss.

In order to put more emphasis on the small values, i.e the earliest activation times of the normalised propagation map A , we reverse the order of the values by using $2M - A$ as input. We also only calculate the reconstruction error in the myocardial binary mask M , named RMSE_M . Therefore we use a weighted RMSE formula RMSE_w defined as follow:

$$\text{RMSE}_w(A) = \text{RMSE}(A, 2M - A) = \sqrt{\text{MSE}(A) \times (2M - A)} \quad (3.3)$$

where: the myocardial binary mask $M \in [0, 1]$; the cardiac activation map normalised $A \in [0, 1]$ and MSE is the Mean Square Error.

There are two terms used in the reconstruction loss. First, we calculate a reconstruction loss between the encoded data from the z -latent vector and the ground truth activation maps L_z , Eq. (3.4), to get a more consistent loss function:

$$L_z = \text{RMSE}_w(x, P(x|z, c)), \text{ where } z = Q(z|x, c) \quad (3.4)$$

Second, we add a reconstruction loss from the multivariate Gaussian $L_{\bar{z}}$, Eq. (3.5), to improve the generative aspect:

$$L_{\bar{z}} = \text{RMSE}_w(x, P(x|\bar{z}, c)), \text{ where } \bar{z} \sim \mathcal{N}(0, 1) \quad (3.5)$$

$P(x|\bar{z}, c)$ can be generated independently from the input activation map, x , while it integrates the information from the conditions c . This can be used to generate activation maps from ECG signals and cardiac masks only.

The total loss function in our model is a sum of the two different RMSE_w and the KL divergence loss $L_{\beta\text{-CVAE}}$. Therefore, the total loss function to minimise when training our model is:

$$L_{ECGI} = L_{\beta\text{-CVAE}} + L_z + L_{\bar{z}} \quad (3.6)$$

3.2.3 Synthetic Data Generation: ECGI Forward Problem

In order to evaluate the proposed methodology, we built a synthetic database from six cardiac anatomies based on CT-scans provided by the CHU Bordeaux, France, see Figure 3.3. All six patients showed post-infarction scar resulting in severe thinning of the left ventricular wall. We generated synthetic electrical data with the forward problem of electrocardiography. In our case, we use a dipole formulation, assuming that the torso domain is homogeneous and infinite [Chávez, 2015].

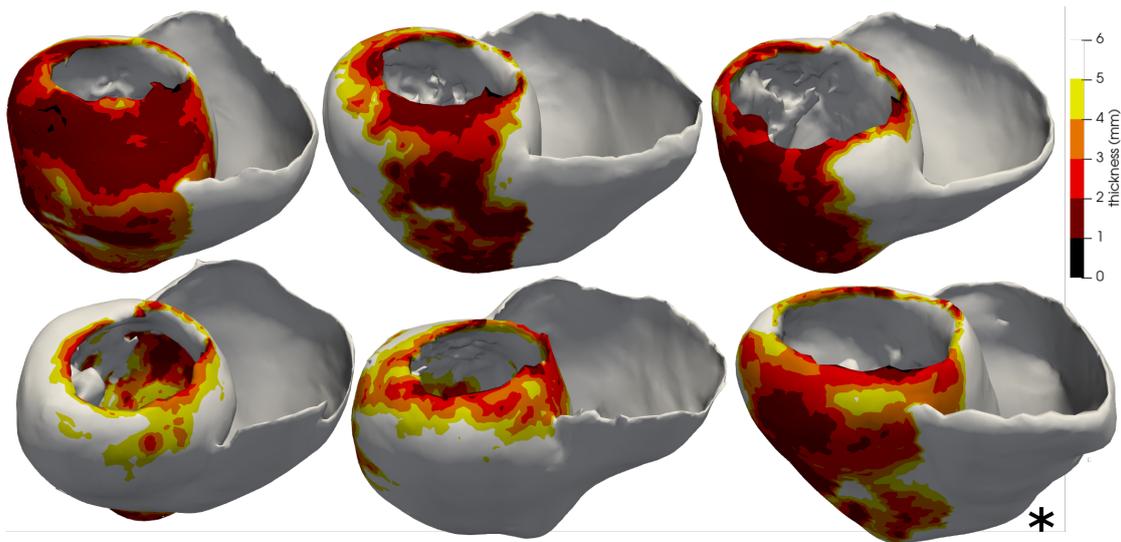


Fig. 3.3: Five cardiac anatomies from CT imaging used for training and one (*) for testing. The thickness maps show the infarct scars.

We proposed a fast method to simulate the activation maps of the heart and their corresponding body surface potentials, Figure 3.4c. It is based on the Eikonal Model on a Cartesian grid from image segmentation [Cedilnik, 2018]. The Eikonal model is a fast generic model of wave front propagation; its results in terms of local activation times are comparable to more sophisticated models [Potse, 2006] and it is efficiently computed with the Fast Marching method (FMM) [Sermesant, 2005]. The Eikonal model takes as

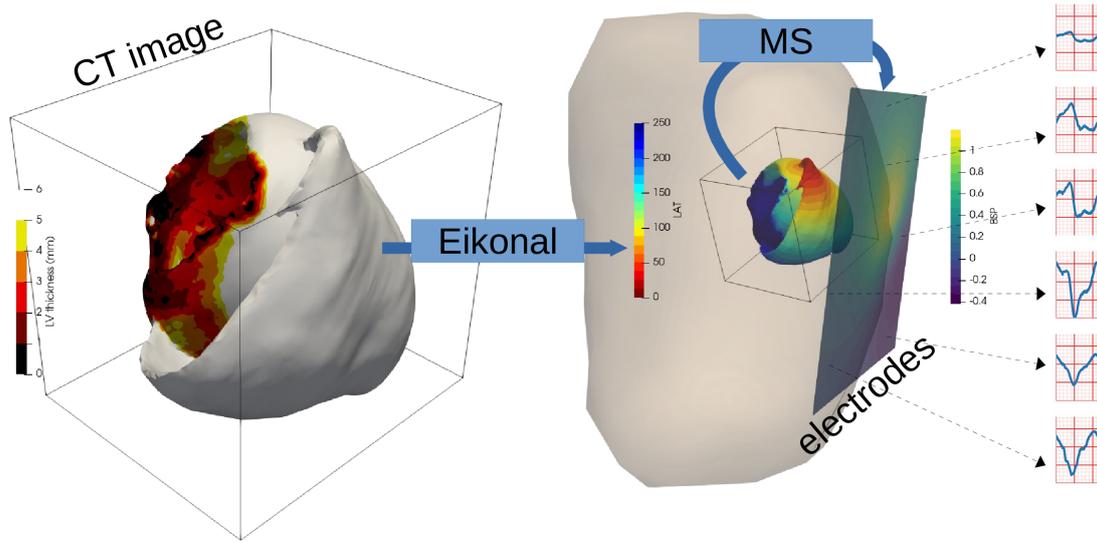


Fig. 3.4: Synthetic data simulation process: myocardial wall thickness from patient image is used to parameterise the Eikonal model that generates activation maps (LAT: local activation time). The Mitchell-Schaeffer model with the dipole formulation are used to generate BSP mapping signals.

input the myocardial wall mask, a local conduction speed s for each voxel v of the given wall mask and the pacing zone. The model equation writes:

$$s(v) \|\nabla T(v)\| = 1 \quad (3.7)$$

with $T(v)$ the local activation time in voxel v .

For these simulations, we vary the local conduction velocity according to the local myocardial wall thickness, as chronic infarcts generate thinner walls with slower conduction velocity; this approach has been shown to efficiently reproduce intra-cardiac activation maps recordings [Cedilnik, 2018]. For each simulation, to initialise the wave front propagation, we chose randomly (from a discrete uniform distribution) a single pacing onset where the heart had a wall thickness superior to 6 mm. We defined a stopping value for the propagation at 250 ms.

To numerically compute the Forward problem, i.e estimate the BSP from the cardiac data, we first model the 100 torso electrodes by a 10×10 Cartesian grid on the front-side of the torso. We consider that the torso-domain is homogeneous and infinite. To generate the BSP we use the dipole formulation from the volume conductor theory where every element of the cardiac volume is a time varying current dipole [Giffard-Roisin, 2017]. Moreover, each activation time obtained with the Eikonal model is associated with an action potential from the cardiac cell Mitchell-Schaeffer model [Mitchell, 2003], which is

a refined method of the Heavyside function sometimes used to link activation time and transmembrane potential. The local current density j_{eq} is written as:

$$j_{eq} = -\sigma_E \cdot \nabla V = -\sigma_E \cdot \frac{\partial V}{\partial t} \frac{\partial t}{\partial x} = -\sigma_E \cdot \frac{\partial V}{\partial t} \nabla T \quad (3.8)$$

with σ_E the local electrical conductivity, ∇V the spatial gradient of the potential V and ∇T the gradient of the activation map given by the Eikonal model [Cedilnik, 2019]. We can therefore directly compute the local dipole from the gradient of the activation map and the time derivative of a transmembrane potential model (Mitchell-Schaeffer).

In total, we simulated 6 000 activation maps and the corresponding BSP using 6 patient image segmentations from CT images. The choices of 120 simulated data in 2D [Bacoyannis, 2019] and 6 000 in 3D were done empirically taking into account that 3D extension of the 2D model requires more training data, as we have more hyperparameters to optimise but also considering the trade-off between the computation time as well as the cost of energy consumed during the training phase and the acceptable accuracy of the obtained results. We resampled the activation map and the myocardial wall to have the same dimensions $100 \times 100 \times 100$. The signal input is kept at $10 \times 10 \times 250$ (10×10 correspond to the grid's dimension and the third dimension to the propagation's time $250ms$).

3.2.4 Computational Details

The model was trained on 5 patient geometries, see Figure 3.4a, including a total of 5 000 activation maps and their corresponding BSP, which were divided into 8:2 ratio for training and validation datasets. We then tested the trained model on 200 activation maps and their corresponding BSP from a new geometry unseen during the training Figure 3.4b. The neural network design, training and testing were implemented using Keras API of Tensorflow 2.0 ².

To adapt the signal shape, we used two 3×3 2D convolutional layers with the number of strides of 2 and 5 and the number of output features of 200 and 100, respectively. The encoder of our network consists of two successive stride 2 convolutional layers (Figure 3.2), one flatten layer followed by one dense layer. Each layer has a $3 \times 3 \times 3$ kernel and all the activation functions are \tanh . The bottleneck layers (μ , σ , z) are fully-connected. From different tests, the latent code size was set to 25. The decoder has a fully-connected layer which is then reshaped into a 3D matrix before applying a succession of transpose and normal convolutional layers. The model was optimised using the total loss described in Eq. 3.6, without any additional weight applied to the loss terms. The best value for β , the penalisation term in the loss function, was 10. We

²https://www.tensorflow.org/api_docs/python/tf/keras

used Adam [Kingma, 2014a] optimiser with a learning rate of 10^{-4} , with a batch size of one for 240 epochs. The computation time for training was 64 hours on a NVidia Tesla V100 (32GB of RAM).

3.3 Results

When testing the method, our probabilistic generative model allows us to generate multiple solutions from a single input, by sampling the latent space. To compare our predictions with the ground truth simulated activation maps, we generated 10 different activation maps for each case and computed the mean and standard deviation of these predictions. Figure 3.5 shows a 3D visualisation and a slice of the estimated mean activation map over the 10 predictions (b) in comparison with the ground-truth simulated cardiac electrical activity (a). (c) represents the standard deviation map for these 10 predictions, and (d) the error map between the generated and the ground truth maps.

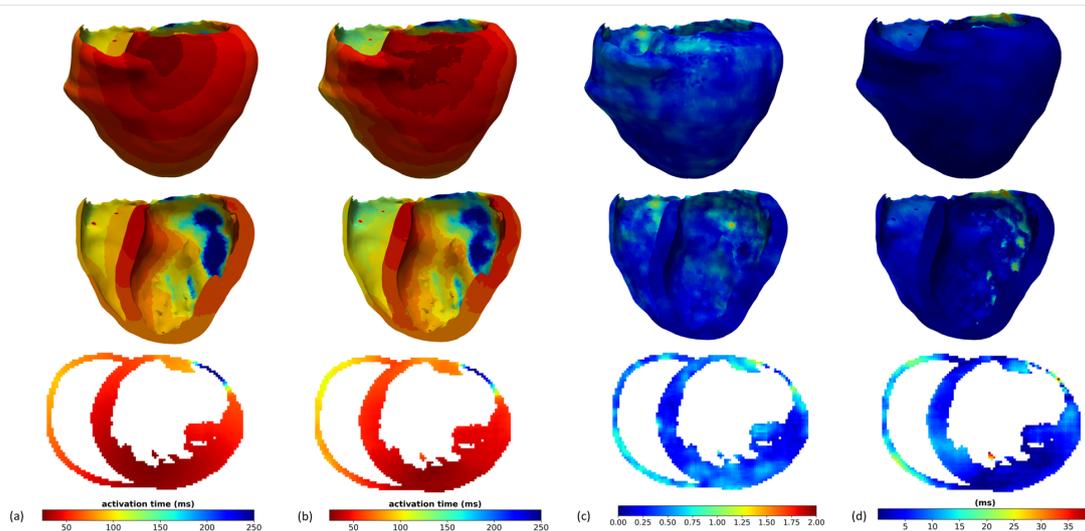


Fig. 3.5: (a): Ground Truth activation map, (b): Activation map generated by the presented method, (c): standard deviation map calculated over 10 predictions, (d): error-map, difference between predicted and simulated activation maps. All results are in 3D, showing the middle x-axis slice in the bottom of the figure.

We evaluated our model by testing it on a new patient anatomy over 200 activation maps. Due to the probabilistic property of our method, we predicted 10 distinct activation maps for each BSP data.

The proposed method is able to generate volumetric activation maps with a good accuracy on this simulated data: the mean absolute error is 11.33 ms with 4.10 ms standard deviation on this testing set.

Small errors and small standard deviation in Figure 3.5c and Figure 3.5d indicate that the reconstruction performs well and in a consistent way, while large values mean that the reconstruction is suffering. The zones with the highest error correspond partly with the zones with the highest standard deviation in the maps. This underlines the potential use of our probabilistic generative model to help in quantifying the uncertainty in the predictions.

From the representative results shown in Figure 3.5 we can observe that the weight applied to the loss function influences the performance of the reconstruction in the later regions of the wave propagation. In opposition, zones where the propagation starts are better reconstructed.

3.4 Discussion

The proposed approach demonstrated great accuracy on this synthetic database with a good robustness to anatomical variations. However the goal of this work is to apply it clinically for diagnosis and therapy planning. So it is necessary to detail the limitations we have at this present stage in order to consider the required extensions.

To generate our database, we only considered single pacing sites. A first extension to our work will be to simulate data with multiple onsets. This is straightforward to implement, but the accuracy of the results has to be evaluated.

Second, we use the Eikonal model, which is a simplified model in terms of transmembrane potential capabilities in comparison to more complex models as the Bidomain one. However, we believe that this model can be used in many clinical cases. Indeed, the spatial dynamics of the Eikonal model and the reaction-diffusion equations are extremely similar [Wallman, 2012]. The local conduction velocity modifications can represent the impact of infarcts on activation maps [Cedilnik, 2018]. Thus, we believe that this fast method to simulate activation maps is relevant in a context of myocardial infarction. The limit of this model is reached for arrhythmias or pathologies modifying local transmembrane potential profile. From a more global and mathematical point of view, the Eikonal model provides a very strong prior on spatiotemporal correlations for solving the ECGI inverse problem.

Moreover, our method makes the hypothesis about isotropic propagation in the heart. Although heart anisotropy may not play an important role for ECGI because of the spatial scale at which it is working, we know that it has an important impact on local propagation front and thus in the forward problem. In the future, we are going to include the fibres in our forward problem formulation, which will be an easy extension.

In order to evaluate our model, we generated 10 activation maps for each case. It would be interesting for each one of these cases to vary the number of activation maps we generated and to estimate the effects on the accuracy of the model and the uncertainty quantification.

Moreover, this method has to be evaluated on real data. The differences between the training data and the testing data can severely impact performance. Robustness to signal noise and missing leads, a classical problem in ECGI, has still to be quantified. We believe that adding noise to our simulated database should help. We also intend to train the model on a wider variety of simulated data and also on some real data with partial ground-truth from catheter mapping. To help moving from synthetic to real data, we may use transfer learning [Giffard-Roisin, 2018].

3.5 Conclusion

Our proposed β -CVAE is able to generate a 3D volumetric map of the cardiac electrical activity for a new geometry from non-invasive BSP and cardiac imaging data. During the training, our model is designed to learn spatiotemporal as well as geometrical information. This novel formulation enables to directly introduce all the imaging information on cardiac substrate within the ECGI methodology and therefore fully exploits the available patient information.

This is the first time an ECGI method is able to reconstruct the activation pattern throughout the myocardium in a volumetric manner at such resolution, while integrating 3D substrate information from imaging and within a probabilistic framework. This promising formulation now has to be evaluated on real clinical data.

Fusion of Body Surface Potential Mapping and Cardiac Imaging for ECGI Evaluation

Contents

4.1	Introduction	49
4.1.1	Clinical database	49
4.1.2	Motivation	50
4.1.3	Current data status and brief overview of the strategies employed	50
4.1.4	Material	54
4.2	Clinical Data processing	54
4.2.1	Extract the data	54
4.2.2	Segmentation	54
4.2.3	Cardiac registration	56
4.2.4	Generic Torso mesh	57
4.2.5	Cardiac mesh	61
4.2.6	Transfer the invasive measurements to the optimal mesh . . .	65
4.3	Conclusion	80

In this chapter we describe the processing framework that we developed in order to fuse clinical BSP mapping data with associated CT-derived meshes for ECGI evaluation. First, we explain the purpose of having "usable" clinical data in our study, then we describe the different types of data we use and finally the different steps we implemented in the processing pipeline.

4.1 Introduction

4.1.1 Clinical database

The data sets used in this work were provided by the CHU Bordeaux, and were shared between the Inria teams of Bordeaux and Sophia Antipolis. The original database consisted of 15 patients, from which 11 patients were finally included in our study. For each patient retained for this study, we collected three components separately:

- Computed Tomography data (i.e., the cardiac and torso CT images).
- CARTO[®] EAM mapping data: containing Matlab files where the geometrical data (i.e., the coordinates of each recorded point on the epicardium, and on the RV and LV, respectively) along with the values of the catheter-based recordings of the potentials on the heart are stored. Three type of voltages are available: unipolar, bipolar and bipolar lat (i.e., local activation times).
- ECGI data: this is raw data which contains: the geometry of the torso electrode vest, the heart, the aorta, as well as the BSP values and the potentials on the heart reconstructed by the CardioInsight technology.

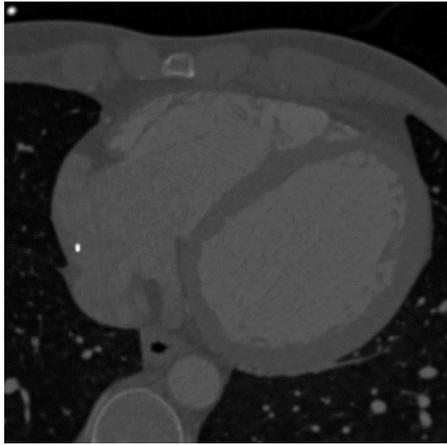
4.1.2 Motivation

The direct aim of this section is to obtain a set of realistic data on which we may evaluate our model for solving the ECGI problem. The processing of real clinical data, especially CARTO[®] EAM data obtained invasively, could also be seen as ill-posed in the sense that there is no unique solution, nor there is a clear and established process. For the majority of the cases analysed, we find out, only post analysis (when we are confronted with an inconsistency in the data or a problem), the need for additional data processing. In this framework, we will detail our data processing steps and the methodologies we have applied. Thus, implicitly, we hope that this chapter will provide ideas, facilitate and accelerate the implementation of processing, and help the clinical data processing community.

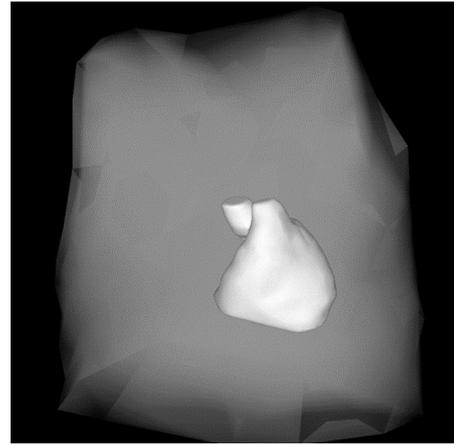
4.1.3 Current data status and brief overview of the strategies employed

At this stage, we have clinical data from three different types of modalities, namely: EAM, CT-scan and ECGI. We could categorise our data as follows: i) on the one hand, we deal with anatomical data including the heart as a whole, along with the torso and the multi-electrode vest for torso; and ii) on the other hand, potential measurements recorded on the torso as well as on the heart.

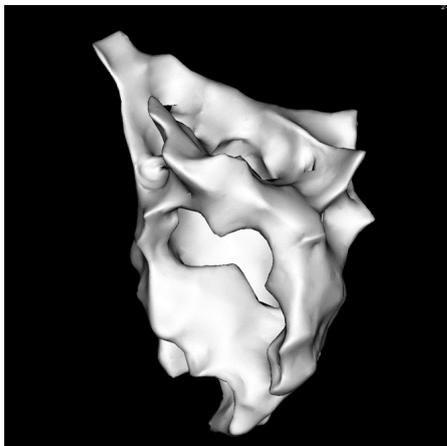
The anatomical data are from CT-scan images and EAM. From the CT-scan images we are able to obtain the geometry of the heart (see an example in Figure 4.6d) as well as the torso vest geometry (as illustrated in Figure 4.4). To do this, the main processing used is segmentation, which is described in section 4.2.2. From the CARTO[®] data of the EAM, we have independent access to the geometries of: the epicardium; the RV; and, the LV. For the three previous geometries, we also have recorded values of the bipolar



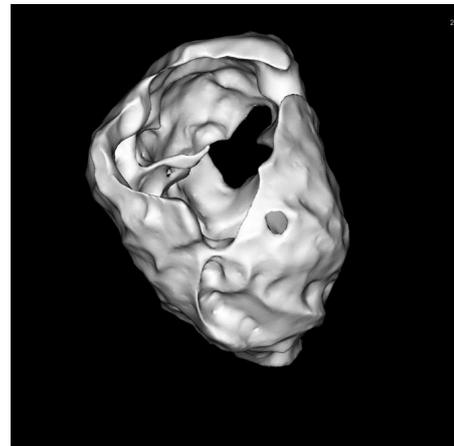
(a) CT-scan image visualised with MUSIC[®].



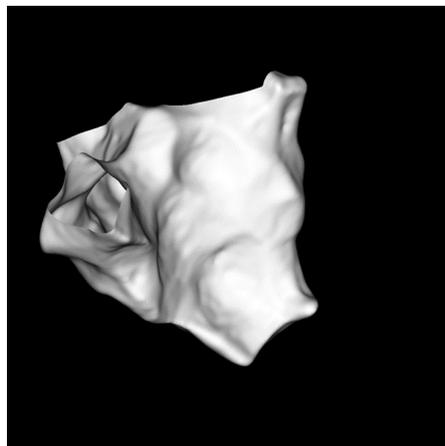
(b) Meshes of the torso, the aorta and the pericardium extracted from the raw data of ECGI



(c) LV



(d) Epicardium



(e) RV reconstructed from CARTO[®] data

Fig. 4.1: Example result from the data extraction process: a) original CT image of the heart, (b) reconstructed torso, aorta and pericardium meshes from ECGI raw data; (c) (d) and (e) respectively, are the reconstructed LV, epicardium and RV surfaces from the CARTO[®] data.

lat, unipolar and bipolar voltages resulted from the EAM procedure. However, for the reasons discussed in section 1.2.2.2, we do not have voltages for all the myocardial points. This results in areas without images, manifesting as holes in our surface meshes (see an example Figure 4.1c). This is a problem that needs to be solved, which we address in section 4.2.6.1. Regarding the BSPs, we have all the signals at each of the 252 electrodes of the CardioInsight vest. Figure 4.5d shows an example where the BSPs of a patient are displayed on the corresponding torso mesh.

The data we have at our disposal and which we have detailed before, are necessary but not sufficient. Indeed, in order to be able to evaluate our model based on Deep Learning with clinical data, we need a larger amount of data in order to train our model. Thus, the solution we have adopted is to create a synthetic database including activation maps and BSPs from the geometries of the heart and of the torso electrode vest, which have been processed and are derived from the clinical database. The cardiac geometries on which we simulated activation maps, are the meshes resulting from the segmentation of the CT-scan images. Further processing of the cardiac surface meshes obtained by segmentation is necessary so that they have, notably, the attributes that our simulation pipeline requires as input. These processing steps are discussed in section 4.2.5.

In section 3, where we trained and tested our DL based model exclusively with synthetic data, the electrode vest was modelled by a Cartesian grid of 10 by 10 electrodes. In the current context, where we have the clinical data: recorded BSP and electrode vest segmentations, for each of the patients in the database, we have refined our electrode vest modelling. The electrode vest is now modelled by two parallel Cartesian grids each containing 10×10 electrodes. Figure 4.4 illustrates our proposed model of the electrode vest. Our method for customising each of the torso models to that of a patient's torso electrode vest is described in section 4.2.4.

Although the final meshes obtained for the heart and torso come originally from the same modality, i.e. CT-scan imaging, they are not segmented from the same image acquired at the same time. This implies that the heart and the modelled electrodes vest are not in the same spatial reference frame. This problem to be solved is not unique to the geometrical data to be used for the simulation database. It is common to the data that constitute the test database. Indeed, our multi-modal data for a single patient are all acquired at different times and therefore none are in the same spatial reference frame. At this stage, we have chosen to consider that the fixed spatial reference frame will be the one of the cardiac images obtained by the CT-scan. Thus, the geometries of the RV, the LV and the epicardium from the CARTO data are registered to the corresponding CT-scan data (see section 4.2.3). The electrode vest is registered to the patient's torso segmentation, by applying the transfer matrix resulting from the registration of the epicardium of ECGI data to the CT-scan epicardium.

However, even if we have now all the data needed (i.e., cleaned CARTO EAM meshes, cardiac mesh model and torso vest of electrode model with BSP) to built our test dataset, we have several processing steps to carry out to achieve this. Since the geometries of the CARTO data and the CT-scan segmentations do not resemble each other very close, it was necessary to develop suitable projection methods for these data. Thus, for each of the cardiac geometries (i.e., LV, RV and epicardium) we propose a different method to project the activation times on the corresponding geometry from the CT-scan segmentation. The methods are described in section 4.2.6.2. The evaluation of the three different methods presented in section 4.2.6.3 confirms that each of the proposed projection methods is more suitable for a given geometry. Before the surfacic activation times of the three geometries (LV, RV and epicardium) can be integrated on the surface of the same cardiac mesh representing the entire heart, it is necessary to perform another step that consists on synchronising the activation maps of the LV, RV and epicardium. It should be noted that we encountered difficulties with this particular step. Specifically, we have managed to synchronise only some of these maps but not all, leaving us with open questions about how to resolve this issue. All cases are listed in section 4.2.6.4. Finally, for the cases where we managed to synchronise the maps, using the labels of the cardiac mesh, for each point of the cardiac mesh we detected its matching point in the set of LV, RV and epicardial synchronised maps and we associated the value of the bipolar voltage to it. Thus, we were able to supplement the test database, which already consisted of the BSPs on our electrode vest model, with the clinical activation times on the cardiac meshes that we modelled.

For the sake of clarity, since we handle many meshes in this chapter, let us define the most important of them:

- Let: LV_C , RV_C and Epi_C be respectively the LV, the RV and the epicardial meshes from CARTO EAM data, with the Bipolar lat values.
- Let: LV_s , RV_s and Epi_s be respectively the LV, the RV and the epicardial meshes segmented in the cardiac CT-scan image.
- Let C_m be the final mesh obtained after all the successive processings starting on the original cardiac CT-scan image. This is the main cardiac mesh obtained at the end of section 4.2.5 and will be used to project the activation times measured invasively in the clinic, and also as an input mesh for our simulation pipeline.
- Let T_s be the torso vest of electrodes obtained by segmenting the CardioInsight vest in the CT-scan images.
- Let T_m be the generic torso mesh that we modelled and proposed (in this chapter) in section 4.2.4.

4.1.4 Material

- Visualization ToolKit (VTK) [Schroeder, 2006]: is an open source object oriented software that supports a wide variety of algorithms for three dimensional computer graphics, modelling, visualisation. It is widely used by researchers, particularly in the medical field, for visualisation but also for image processing.
- MUSIC[®] (Multi-modality Platform for Specific Imaging in Cardiology)¹: is a proprietary software based on the open-source software MedInria and has been jointly developed by Inria and the Liryc Institute. Besides the fact that this software interface offers a very good visualisation of the data and this for different formats (.mha, .vtk, .xml,...), it also has the advantage of proposing a collection of diverse tools for processing cardiac images.

4.2 Clinical Data processing

4.2.1 Extract the data

The first step was to extract the necessary information from the clinical data and export it in a format that can be easily manipulated and visualised. From the raw ECGI data, the Carmen Inria team extracted and processed the signals that were shared with us in Matlab format and the geometries in VTK files. Both teams worked on the cardiac segmentation. The RV, LV as well as the epicardium were segmented from CT images. Finally, the segmentations obtained in Bordeaux were used, and only the corresponding segmentation methods developed at INRIA Bordeaux are detailed in section 4.2.2. The extractions from the Matlab files containing the invasive mapping data were processed jointly by the two teams. Vtk meshes were created containing the values of the unipolar, bipolar and lat bipolar EGMs.

4.2.2 Segmentation

Cardiac segmentation

The first step consisted in using an automated segmentation method based on deep learning [Cedilnik, 2019]. This method based on a previous work [Jia, 2018], allowed us to obtain a segmentation of the LA, endocardial LV, LV wall and epicardial RV. The final masks of the LVA, epicardium and endocardial LV are obtained after applying a succession of image processing operations (closing, median filter, dilation). The endocardial RV

¹<https://www.ihu-liryc.fr/en/music/>

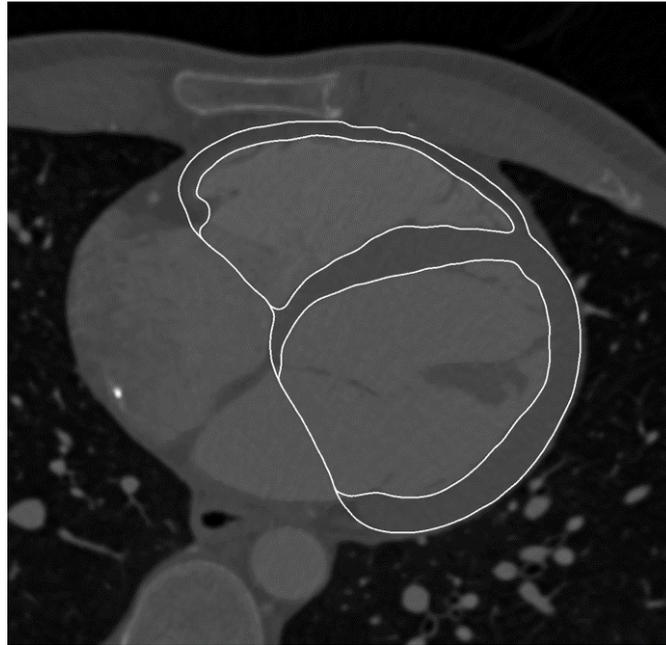


Fig. 4.2: Cardiac CT image in which the solid white line represents the contour of the segmented endocardial and epicardial surfaces.

was obtained by erosion of the epicardial RV. However, this mask contained unnecessary details for us (such as the papillary muscles) which were eliminated by a succession of operations similar to those mentioned above. Finally, a mask joining the endocardial RV and LV and the LVA was created. In the next step, a VTK filter was used to create an epicardial mesh and an endocardial mesh, respectively. Since the walls of the endocardium and epicardium are often very thin, this results in the endocardium mesh overlapping the epicardium mesh. In order to solve this problem and force the endocardium to retract, a binary operation was first applied followed by a transformation. Lastly, this mesh was remeshed with MMG [Dapogny, 2014].

Torso segmentation

The segmentation method used for the torso is based on the seed point methodology. This method extracts all connected pixels that have their intensity within a specified range. The image binary scale is inverted to assign the torso pixels a value of 1 and the rest are assigned a 0 value. An erosion operation followed by a dilation on the mask allowed us to apply a smoothing step and also the CardioInsight vest to be eliminated. An initial mesh was obtained by applying a VTK filter, and further remeshed with MMG.

4.2.3 Cardiac registration

We acknowledge that at this point we have several cardiac meshes (from CT-scan, invasive imaging and ECGI data) for each patient, but these are all obtained from different modalities and were not acquired at the same time. Both the orientation of the heart and its position in space play an important role. Therefore, in this context, we had to register these different meshes such that they were in the same space and orientation. In other words, we tried to superimpose these meshes as accurately as possible. Before being able to register them, a reference frame must be defined. For simplicity, we chose: Epi_s , LV_s and RV_s (the meshes from the CT-scan images) to be this reference. Thus, in the following steps, C_m will be the mesh used by our Deep Learning algorithm for training. The objective of the model will be to reconstruct the activation map of a new cardiac geometry that was obtained from the CT images instead of the invasive imaging. Moreover, it is the least complex and most complete geometry among all our data. This was the best option, given that we have to fit meshes from two different modalities to a reference mesh. Notably, as we are dealing with very complex geometries, it was not possible to use a fully automatic registration method to perform this step. Therefore, we registered the meshes manually via MUSIC[®]. More precisely, we used a function that allowed us to perform the registration by landmarks. For this purpose, similar points on the two meshes of the same geometry (e.g. on LV_i_C and on LV_s) were first annotated, as illustrated in Figure 4.3, then an affine transformation was applied and a transformation matrix was computed. To register T_s , a suitable solution was to apply to it the transformation matrix which was obtained from the registration of the epicardium from ECGI data on Epi_C . As the 3D surface meshes of the epicardium, the LV and the RV from the CARTO data are independant, the registration was done independently for Epi_C , LV_C and RV_C , respectively.

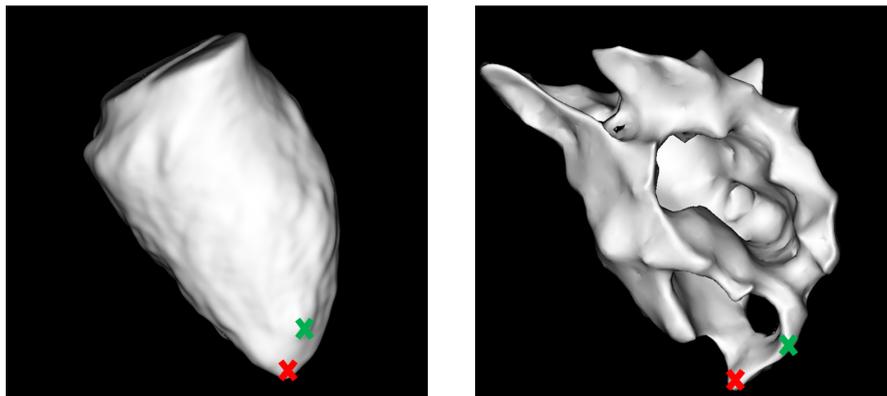


Fig. 4.3: (Left) the reference mesh which is the LV from CT-scan segmentation where the crosses represent the annotated points, (Right) LV from CARTO[®] data with two annotated points corresponding to the ones from the left.

Our aim is to evaluate our model to solve the inverse problem with real data. However, in practice it is impossible to use only clinical data to do this. Deep Learning models require a large amount of data to be trained and currently there is no available clinical database for ECGI meeting such criteria. Therefore, we need to create a large synthetic database from our small database of 11 patients. This 'simulated' database corresponds to the forward problem of the ECGI. We must therefore first decide on how we are going to model the patient's torso electrode vest T_m for which the BSPs will be computed. Hence, a good compromise has to be found between making the modelling as representative as possible of what is really happening (of the electrophysiological phenomena and the measurement device) and the computation time for these simulations. To recapitulate, our method requires the following as training data: cardiac geometries; its corresponding BSPs; and, its corresponding cardiac activation maps. What we now have is: LV_s , RV_s and Epi_s , and activation maps but on LV_C , RV_C and Epi_C (meshes from CARTO[®]). The next task is to reconstruct C_m the whole geometry of the heart from the segmentation of different cardiac structures (i.e., LV_s , RV_s and Epi_s). Moreover, this mesh will be used to calculate the associated activation maps and will have to be formatted as required by the simulation pipeline. More details are presented in section 4.2.5. Finally, once trained, the model will be evaluated with real data. In order to compare the activation maps we have reconstructed with the ground truth, we need to transfer the values of the invasive measurements we have on: LV_C , RV_C and Epi_C to the corresponding remeshed mesh C_m) from CT-scan mask.

4.2.4 Generic Torso mesh

In [Bacoyannis, 2021], we modelled 100 torso electrodes by a 10×10 Cartesian grid on the front-side of the torso. Here, we expect our model to have less ease in estimating activation maps for real data. We therefore propose to make the modelling of the electrode vest more complex. In theory, the same electrode device was used for all patients. The CardioInsight vest consists of 252 electrodes. One would therefore expect that 252 electrodes would be activated during all recordings. However, this depends on the electrode adhesion to the skin, and also it is possible that in some cases not all electrodes work. Unlike the conventional 3-12 leads for ECG recording where the cardiologist manually places those on the patient's torso, in the case of the vest the multi-electrodes are attached. Thus, one can assume that the distance between two given electrodes would be unchanged from one patient to another; however, this assumption is not valid since each patient has a different body shape and torso geometry. Therefore, we offer a customised torso model T_m for each patient in the database.

As a first step, we investigated the following method. We calculated the mean distance between two consecutive electrodes on the whole database, and obtained a value of 50mm. This approach has many limitations. A major problem is that in trying to create

spatially 'fixed' grids from 'fixed' inter-electrode distances, the grids are not in the same coordinate frame as the torsos T_s that we previously registered. It would therefore be necessary to carry out a new registration, which will likely induce cumulative errors. Moreover, here we have considered the two grids independently; thus, when a grid electrode has its associated reference mesh electrode, this reference electrode is often associated with electrodes from both grids.

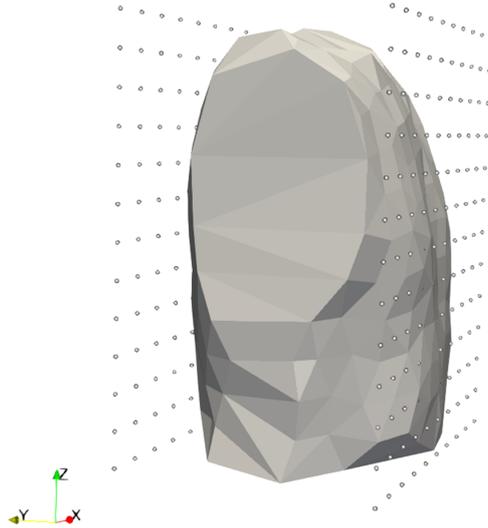


Fig. 4.4: Case of a torso mesh T_s and the electrode grids T_m corresponding to this geometry.

Finally, we model 260 torso electrodes by two 10×13 Cartesian grids (see Figure 4.4). These two grids are parallel to each other, one representing the front of the torso, the other its back, respectively. We chose 260 electrodes to be as close as possible to the 252 of the clinical vest. Each of our Cartesian grids consists of 10×13 electrodes because we took into account that the CardioInsight vest contains more electrodes along the z-axis than along the x-axis. The process can be decomposed into several steps. Below we detail the method for a torso geometry:

- From the registered torso mesh T_s , the coordinates of each electrode and its associated ID are extracted.
- Several parameters are calculated, such as: minimum and maximum positions along the x, y and z axes, torso height and width. The vectors $(\vec{0x}, \vec{0y}, \vec{0z})$ form an orthonormal Cartesian coordinate system with direct orientation.
- The two 2D grids are created from the previous measurements. Ymin and Ymax allow us to fix the position of the front and the back grid respectively. The number of electrodes along the x and z axes is fixed. Xmin, Xmax, Zmin and Zmax are used to modulate the inter-electrode distance along their respective axis.

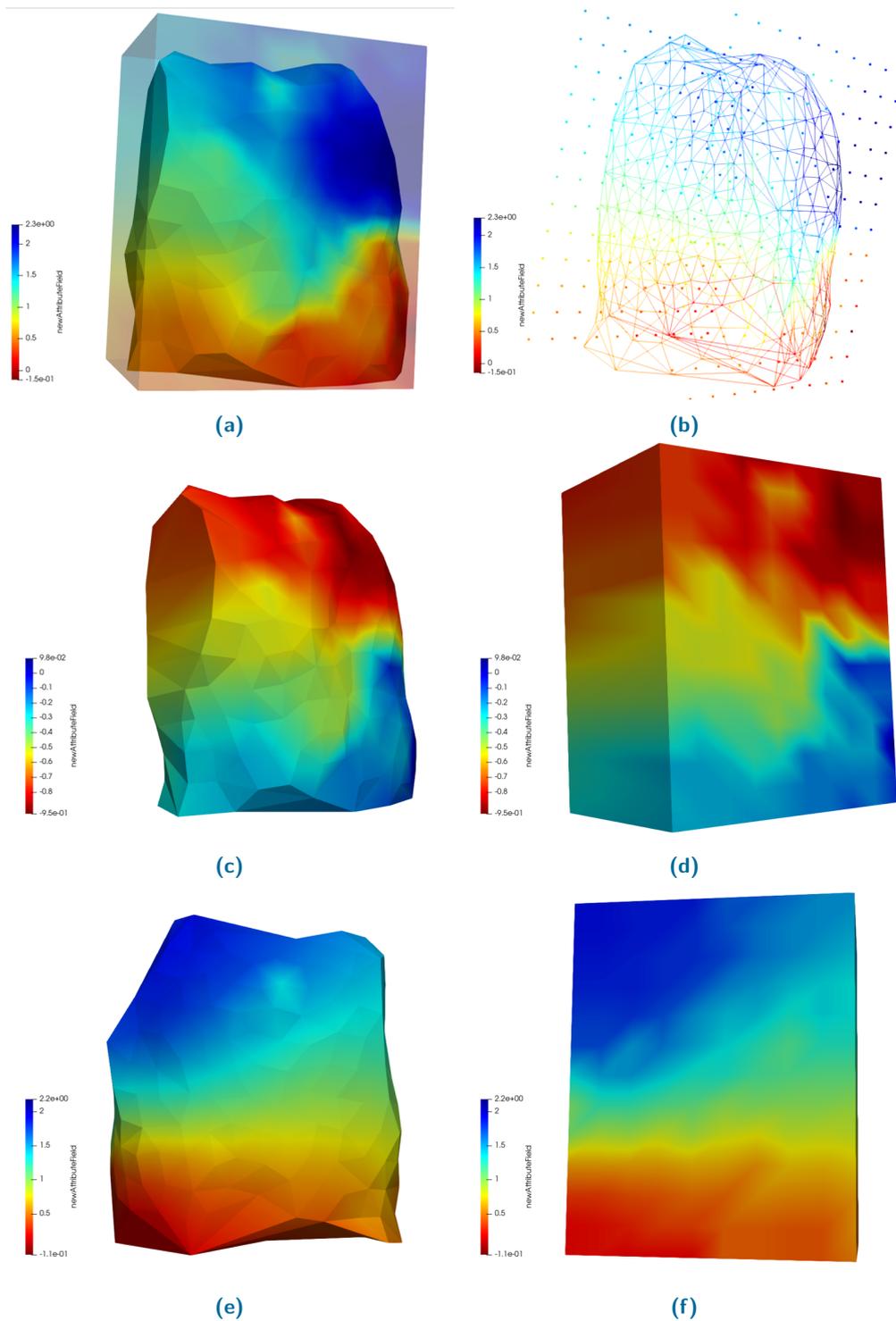


Fig. 4.5: Representation of the Body Surface potentials on the original torso mesh T_s and on the corresponding mesh T_m of the modelled electrode grids: (a) shows the BSP on both: the torso T_s and the electrodes grids T_m , (b) illustrates (a) but representing the electrodes by dots and the torso by wireframe, (c) and (e) show the BSP at a different time step on the clinical torso mesh T_s , respectively in a frontal and back view. (d) and (f) correspond to (c), (e) BSP on the proposed electrodes mesh T_m . The values that appear on the side faces (left and right) of the T_m and T_s meshes (e.g. the righthand faces we can observe in (c) and (d)) are interpolated values that are generated when we represent the surface meshes. These values are never used.

These two electrode grids will be used to simulate the BSP. We will also use them to create the test dataset. To do this, we need to associate each electrode in this mesh T_m with a potential value that we have in the ECGI data (.mat file). The approach adopted is described below:

- For each mesh, we create a matrix containing the coordinates and the IDs of the electrodes,
- From the coordinate matrices we calculate for each electrode of the modelled mesh T_m its nearest neighbour in the clinical mesh T_s . We keep the matching ID, i.e., the ID of the electrode from T_s corresponding to the nearest neighbour found.
- The ID of each points in T_m is replaced by its matching ID,
- The signals in the Matlab file are represented as follows: each row represents a time step and each column represents the signal value at the corresponding point. Thus, each electrode of the mesh is associated with its corresponding signal. This is done for both meshes (T_m and T_s).

Figure 4.5 shows the signal measured on the patient's torso T_s and represented on the corresponding two electrode grids T_m . In Figure 4.5d, we see the side face that connects the two electrode grids. When we represent the BSP on the surface of T_m , the values are interpolated in this area. However, we never use these values.

Sometimes, the electrode assigned as the matching one is found to be so several times, which is expected since we have 260 electrodes on our grids while the clinical data contains 252 electrodes. Note that when this happened too often, we proceeded to solve this problem as follows. For each electrode belonging to the Cartesian grids and whose matching clinical torso mesh electrode is associated with several T_m electrodes, we compute the Euclidean distance between the T_m electrode and the T_s electrode, and saved the ID of the grid electrode with the smallest distance. We then excluded its matching electrode from the database used to find the corresponding electrode and recalculated the matching electrodes for the previous cases. This process was repeated as many times as necessary.

Finally, this processing frame allows us to: confirm that our grids are representative of the BSPs measured in the clinic; prepare the clinical data that will be used for the validation of the Deep Learning approach; and, obtain the meshes on which we will simulate our BSPs that will be used for the training of the algorithm.

4.2.5 Cardiac mesh

The cardiac meshes will be produced (see Figure 4.6d) from the processing of the CT-segmented data. Indeed, the segmentation allowed us to obtain the LV_s , RV_s and Epi_s meshes (as shown in Figure 4.6). The cardiac meshes C_m that we will create will be used: on one hand to have the geometries on which we will be able to simulate the cardiac activation times and thus form a database for the training of our deep learning model; on the other hand to create the test dataset. In the section 4.2.6.2, we will detail the procedure used to create these test datasets. However, before any processing can be performed, it is necessary to define what kind of specific properties the cardiac mesh C_m should have. These properties are related to the choice of the model we will use for our simulations. Here, the mesh must contain the RV and LV labels, as well as the endocardium, myocardium and epicardium labels. An example of a cardiac mesh Figure 4.6d generated from the meshes of different geometries (resulting from the segmentation of the CT-scan), as well as the labels it must contain is illustrated in Figure 4.6. The surface meshes are polydata with triangles. Since the surface meshes contain a very large number of points, the resolution of the Figure 4.6 does not allow them to be represented by their structural elements which are triangles. To have an order of idea, the RV_s which corresponds to the mesh containing the lowest number of points has about 40K.

The steps to generate the mesh are described below:

- LV_s , RV_s and Epi_s meshes were first remeshed using the MMG library [Dapogny, 2014]². The parameters were fixed for all the cardiac geometries of our dataset as follow: hausdorff = 6, hmin = 1.5, hmax = 2.0. The Hausdorff parameter represents the physical distance between the initial mesh and the final mesh. The smaller the value, the more the final mesh will resemble the initial mesh. On the contrary, the larger this value is, the more MMG will allow itself not to respect the shapes of the initial mesh. This can be seen as smoothing. hmin and hmax are respectively the minimum and the maximum size of a triangle edge. This remeshing step is essential to ensure that problems are avoided in the next step.
- We converted the previous remeshed meshes into binary masks. This was done by using MUSIC[®], more precisely the Meshing workspace and the toolbox 'Convert mesh to mask'.

²N.B.: in the following methodology, for practical reasons, if the mesh is currently being processed in the MUSIC interface, we will use the MMG remeshing algorithm implemented in MUSIC by the software designers. Otherwise, for simplicity, we will use ACVD which require limited user interaction, and parameter tuning (only the desired number of points in the remeshed mesh).

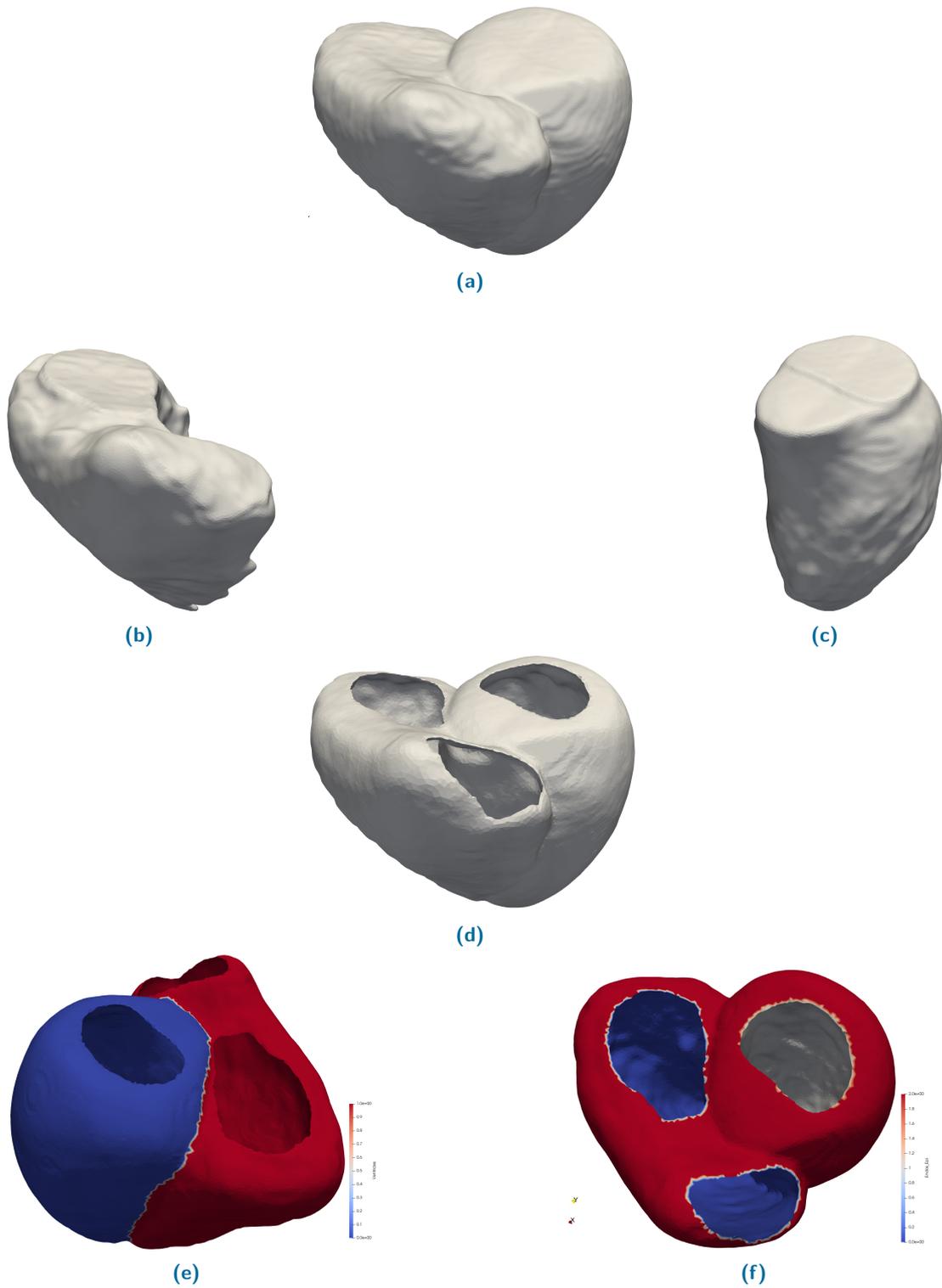


Fig. 4.6: Example of cardiac surface meshes. (a) is the epicardial mesh resulting from CT segmentation (Epi_s , (b) and (c) are RV_s and LV_s also from this segmentation. (d) is the generated cardiac mesh C_m obtained by our processing pipeline, with (e) labelled ventricles and (f) labelled endocardium and epicardium. All the meshes represented here are Polydata vtk with triangles.

- In the MUSIC[®] visualisation workspace, we checked one by one that the created binary masks correspond to their respective geometries on the original scanner image.
- Finally, we obtained the wall mask thanks to the XOR binary operation.
- From the wall mask, we generated the mesh with the Marching cube algorithm [Lorensen, 1987], then this cardiac mesh was remeshed to be uniform.

Label processing:

In order to create the desired labels, we next used the following elements: the binary masks processed above; the meshes of the different geometries (LV_s , RV_s and Epi_s); and, the final generated surface mesh C_m . The process is described below:

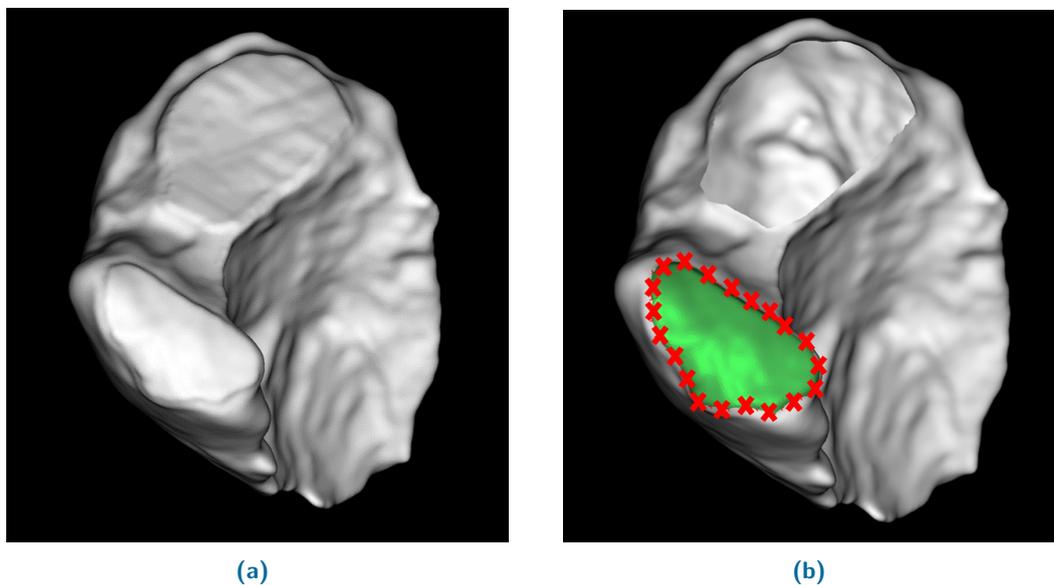


Fig. 4.7: Intermediate processing step for mesh labelling: (a) RV_s mesh from CT-scan segmentation, (b) same mesh as in (a) where one valve has already been removed and the red crosses correspond to the points manually placed to draw the area represented in green which corresponds to the second valve to be removed from the mesh.

- We opened the cardiac valves of the LV_s , RV_s and Epi_s mesh obtained from the segmentation using MUSIC[®]. This was done manually by drawing the outline of the area to be removed from the mesh by placing points one by one. A feature in the software allowed us to remove the triangles associated with this area from the mesh. This procedure is illustrated in Figure 4.7.
- To label the ventricles (LV associated with the value 1 and RV with 0), we used the mesh 'opened' in the previous step and the whole cardiac mesh generated earlier. The general principle is based on the conditions that we set Euclidean distances.

More precisely, for each point of the total mesh C_m , we calculated the closest point of the LV_s mesh without valve and then the Euclidean distance between these two points. Next, we imposed a condition on the value of this distance. Thus, we obtained a mesh where the ventricles are labelled. However, as shown in Figure 4.8a this result is not optimal. Finally, we used the binary masks of the RV_s and LV_s to calculate a distance map according to [Danielsson, 1980], which allowed us to correct the errors (see Figure 4.8b).

- The labelling of the endocardium, myocardium and epicardium, respectively, was done using Euclidean distance. The input data were the meshes from: LV_s , RV_s and Epi_s without the valves, as well as the mesh C_m generated by our processing. No correction is required, as shown in Figure 4.8c
- Once the new mesh was obtained, we cleaned it up by keeping only the necessary arrays (those containing the labels). The generated mesh has about 150 000 points. Then, we uniformly remesh it with less points using the clustering remeshing approach from ACVD library [Valette, 2008; Valette, 2004] which is a simplification algorithm called Approximated Centroidal Voronoi Diagram (ACVD) based on vertex clustering. The clusters are build based on Centroidal Voronoi Diagram. The algorithm returns a uniformly meshed surface. Using the ACVD algorithm, the number of points was reduced to 50 000.

At this stage, we have the necessary mesh for our simulations and therefore for the training dataset. In order to validate our method, we need to have the data measured invasively projected onto the cardiac mesh C_m used by our algorithm. We will next detail this last step of data processing in section 4.2.6

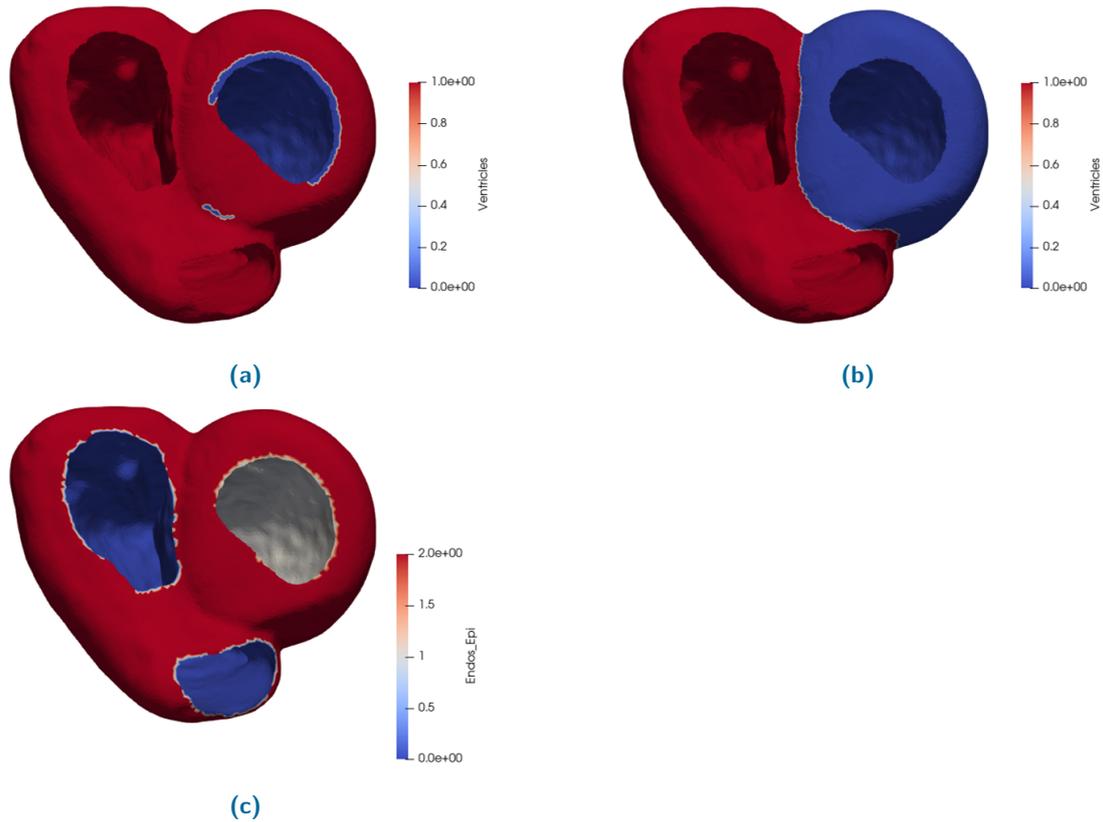


Fig. 4.8: Example of C_m with labelled ventricles: (a) using the method based on the Euclidean distance from the segmented processed mesh data, (b) by correcting the labels with the distance map filtering from binary masks; (c) C_m where the endocardium, myocardium and epicardium are labelled with a method based on the Euclidian distance.

4.2.6 Transfer the invasive measurements to the optimal mesh

To project the invasively measured data (CARTO[®]) onto our cardiac mesh, we proceed in two steps. Firstly we process the (CARTO[®]) data and secondly we develop a method to project the data from one geometry to another.

4.2.6.1 Invasive data processing

The protocol we designed to process the CARTO[®] data is as follows:

- **Extract a surface:** Depending on the patient and geometry, sometimes the mesh contains two surfaces. An example in Figure 4.9 shows that in fact this is the mesh corresponding to the pericardium and not to the epicardium. Indeed, two layers are present, as shown in Figure 4.9b. The inner layer corresponds to the epicardium and what we need. The method for extracting a surface from the mesh is based on the calculation of normals, then the centre of the mesh is calculated and

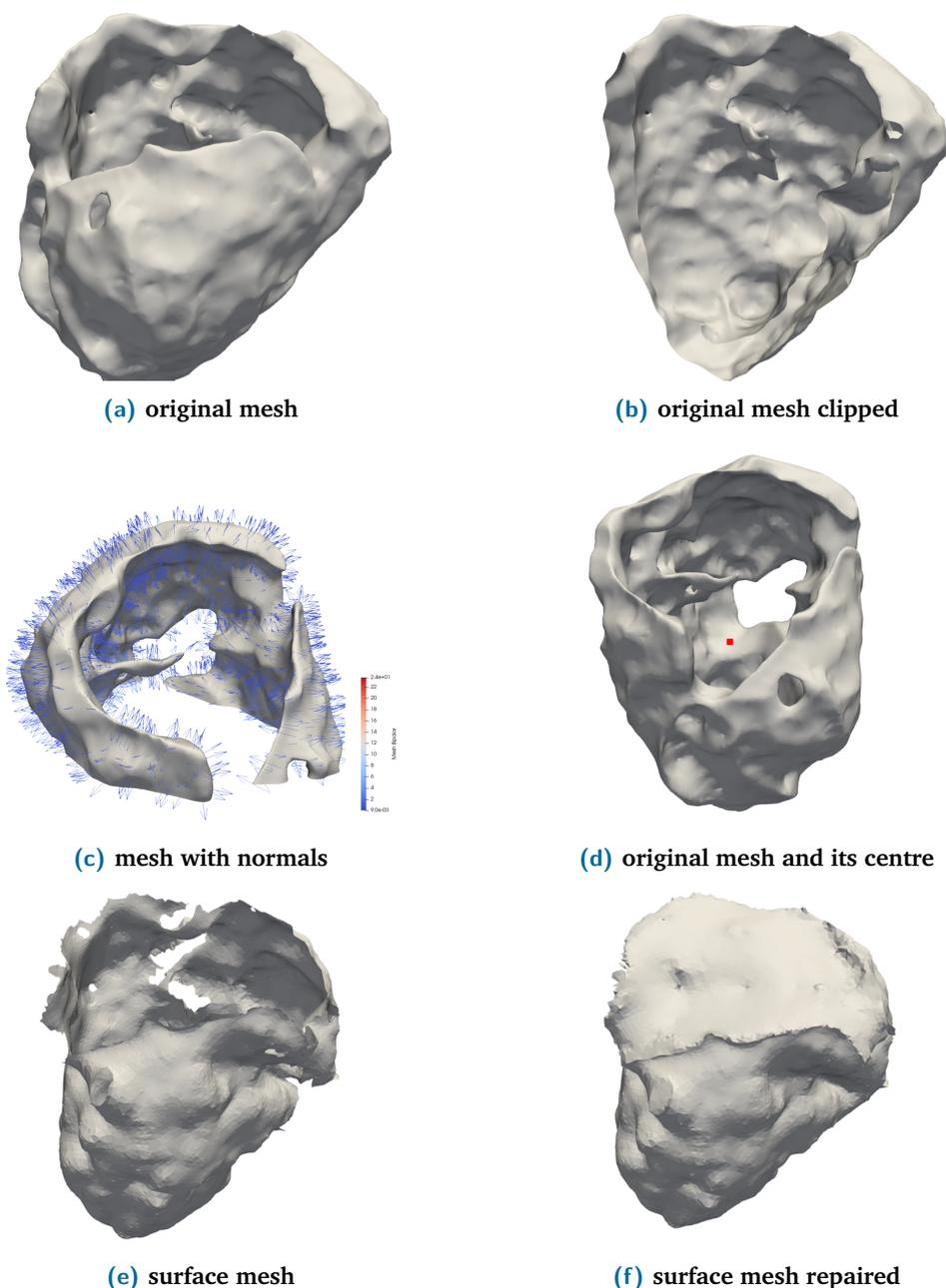


Fig. 4.9: Processing details applied to the epicardium mesh. (a) original CARTO[®] mesh supposed to be the epicardium, (b) view of (a) after clipping, we can see that there is an inner and an outer sheet, (c) blue arrows represent the normals of the points from the clipped mesh, (d) the red dot corresponds to the centre of the whole mesh, (e) is the extracted endocardium surface from (a), (f) represents the surface mesh in (e) after filling the holes.

thanks to these data the orientation for each point of the mesh is calculated. Thus, depending on which surface we want to keep (external or internal), we remove the points whose orientation is positive or negative, respectively. In our case, we keep the points with negative orientation. However, this process creates a problem in the order of indices of the triangles that composed the initial mesh; thus, we need to

implement a method to reorder these indices before creating the epicardial mesh. Figure 4.9 shows the original CARTO[®] mesh clipped for optimal visualisation with the normals to the point represented by arrows and the centre of the whole mesh. It can be observed that the normals of the points towards the inner side point in the direction of the centre, while those on the outer side point towards the opposite direction.

- **Fill the holes:** As mentioned in the section 1.2.2.2, it is impossible during the clinical mapping procedure to record signals from the entire cardiac surface. For this reason one can observe that there are holes in the mesh. This could be fixed next by closing all the holes the vtk meshes. However, this is not straightforward since the closing algorithm leaves some holes unfilled. We initially tested a special VTK filter, `vtkFillHolesFilter`³, which does not close all the holes if the holes are too large or if they have a peculiar shape (e.g. like a figure of eight). The `vtkFillHolesFilter` is used to fill the holes after first detecting the holes. The process is as follows: the edges of the boundary are located, which allows the holes to be identified, then linked together in loops and finally the resulting loops are triangulated. Thus, we also used the repair function provided by the MeshFix[®] software [Attene, 2010]. Meshfix locates problematic regions in triangle meshes, and uses a set of local operations to correct them. This mesh repair automatic method was designed with the aim of modifying the mesh as little as possible. Our experiences have highlighted this objective. Indeed, the meshes repaired with Meshfix are much more regular than those obtained by using `vtkFillHolesFilter`. However, the problem now is that all holes are filled which include the heart valves, as shown in Figure 4.9f.
- **Interpolate the data:** We next interpolate the values of the Bipolar, Unipolar and Bipolar Lat voltages from the inner surface mesh to the mesh without holes. This is done using the 'Data Interpolation' (based on Inverse Distance-Weighted (IDW) interpolation) feature of MUSIC[®].
- **Replace the NaN values:** By interpolating the data from one mesh to another, NaN values are created and then replaced by extreme values (in our case -10 000). This is one possible choice as we could have put the nearest neighbour value, but since the acquired signal values are not precise we tried to avoid the additional erroneous information (which the neural network would learn later).
- **Open the cardiac valves:** This was done manually using MUSIC[®]. This particular step in the framework has been detailed in section 4.2.5.

³<https://vtk.org/doc/nightly/html/classvtkFillHolesFilter.html>

4.2.6.2 Data Projection

We have developed and tested several methods to project the data acquired during the invasive procedure onto the mesh we have generated (see section 4.2.5). Finally, we have selected 3 methods. Depending on the mesh geometry, one method could be more suitable than the other.

The three methods differ only slightly (being extensions of each other), and they are all based on the same general principle proposed by [Cedilnik, 2018].

Method 1:

Here the projection is obtained through a transition to the spherical coordinate system and not from the Cartesian coordinates directly. In order to obtain the spherical coordinates of each point of the mesh, we fit an ellipsoid and defined the new spherical coordinate frame by first calculating the centre of this frame. The main axis of the ellipsoid passes through its centre, the second axis is chosen arbitrarily. Thus, each point P of the mesh is defined by: its distance from the centre of the ellipsoid, and its angles Θ and Φ . One can then compute linear interpolation to a scalar value in the spherical coordinate system (Θ, Φ) from a set of other values (Θ, Φ) . The linear interpolation function chosen was 'LinearNDInterpolator' from the Scipy [Virtanen, 2020] library. In 'LinearNDInterpolator', the process to construct the interpolant involves first triangulating the input data using Qhull⁴ and then on each of the triangles a barycentric linear interpolation is performed.

Method 2:

This is an extension of *Method 1*, incorporating the notion of normals and it follows the same idea as the method presented in section 4.2.6, which was developed to extract a specific surface from the original mesh. The assumption made here is that as the catheter is inserted into either the femoral artery or the femoral vein, it is assumed that the data on the inner surface of the mesh should be the most reliable. Thus, before linearly projecting the values from the source EP mesh (i.e., LV_C , or RV_C , or Epi_C) to the destination generated mesh (i.e., LV_s , or RV_s , or Epi_s), we compute the normal and chose the surface from which the data is to be projected.

Method 3:

This method is an extension of *Method 2*. This method takes into account the fact that the source (i.e., LV_C , or RV_C , or Epi_C) and the target meshes (i.e., LV_s , or RV_s , or Epi_s)

⁴<http://www.qhull.org/>

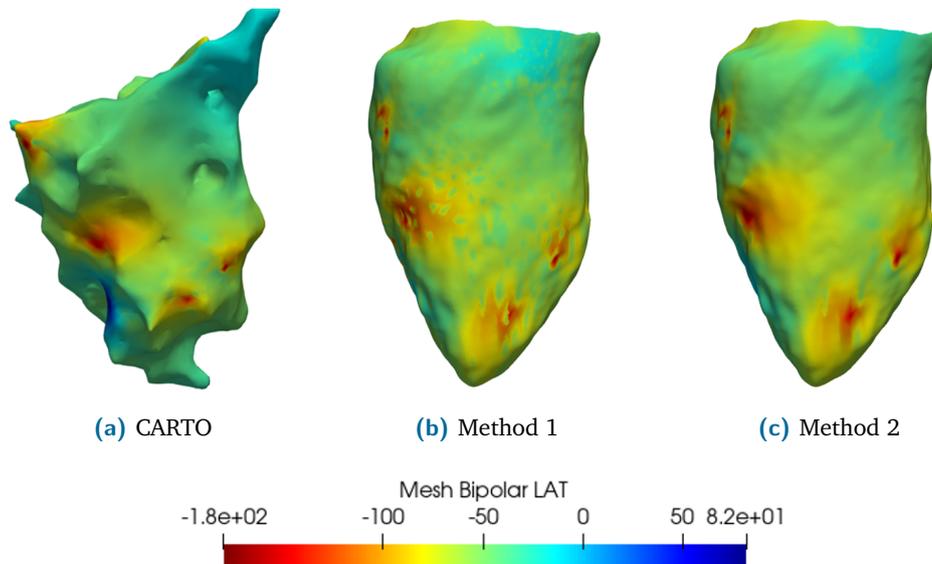


Fig. 4.10: Example illustrating the projection of the bipolar lat onto an LV_s mesh, with the two proposed methods: (a) represents the clinical LV_C mesh with bipolar lat obtained from CARTO navigation system, (b) and (c) are the results of the projection using the mesh presented in (a) as source mesh and applying Method 1 and Method 2, respectively, with a target mesh of LV_s segmented from the cardiac CT images.

are very different. Indeed, they differ by their geometry, but also by their number of points. Here we propose to use an intermediate mesh for the projection step. The target mesh represents the most accurate mesh, but it must be remembered that it has been generated by a process that includes several steps. We propose to first project the values corresponding to the source mesh, to the mesh resulting from the segmentation which does not include a heart valve and which has been remeshed with a uniform remeshing to have fewer points. Then, from this mesh we project the data onto the same mesh but containing more points, and finally we project the values from the latter mesh onto our final mesh C_m .

Methods 1 and 2 were applied to all CARTO[®] (i.e., LV_C , RV_C and Epi_C) meshes at our disposal. Method 3 is dedicated to the epicardial mesh (Epi_C) only. An example showing the process of Method 3 is shown in Figure 4.12. Figure 4.10-4.11 show two cases where Methods 1 and 2 are applied to an LV mesh and an RV mesh, respectively .

4.2.6.3 Assess projection methods

As the geometries of the meshes for which we have clinical measurements are extremely different from the meshes obtained from the CT image segmentation, we cannot evaluate the projection methods other than visually (i.e., qualitatively). This method of evaluation is sometimes complicated. Figure 4.10 shows the projections onto LV_s of the bipolar lat voltages using Method 1 and 2 along with the voltage map obtained from the CARTO

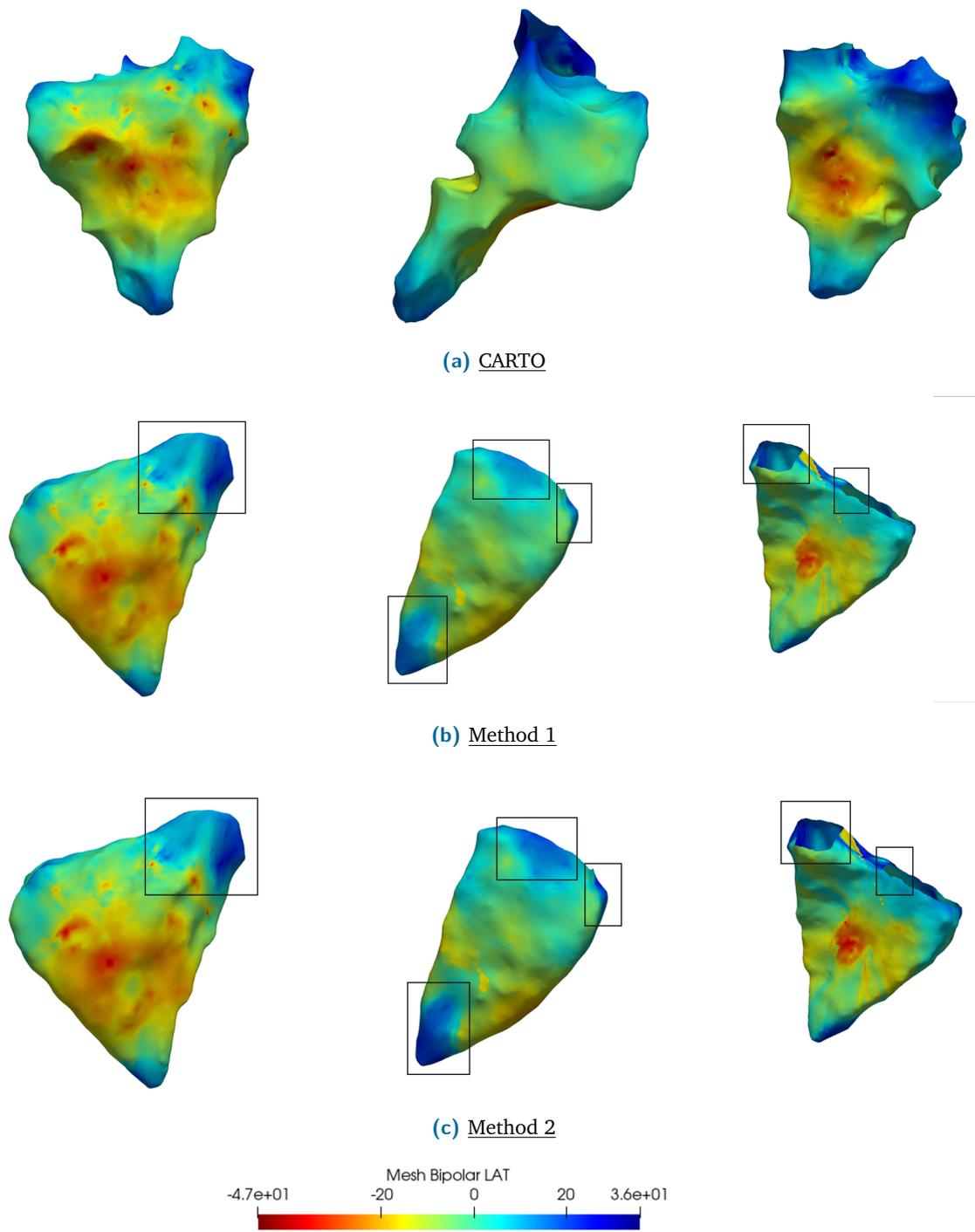


Fig. 4.11: Three different views of the same RV mesh representing: (a) bipolar lat voltage on the mesh obtained from CARTO navigation system, (b) bipolar lat voltage obtained by using (a) with the proposed Method 1 onto RV_s , (c) as (b) but applying the Method 2.

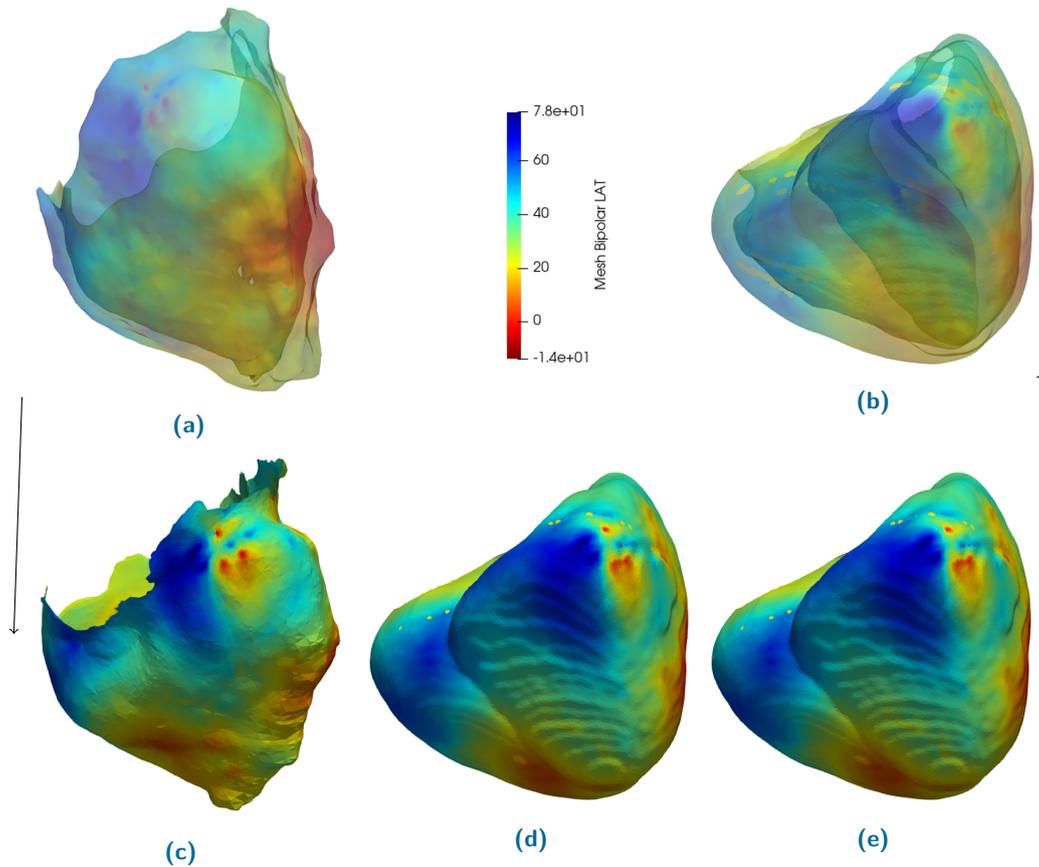


Fig. 4.12: Detailed representation of the Method 3 process used for the projection of (a) local activation times from the CARTO mesh of the epicardium to (b) C_m (the volume mesh used for our simulations). The projection from (a) to (b) is done via intermediate steps: first the LAT points are projected onto (c) the surface extracted from the reference mesh Epi_C , then onto (d) the mesh of the segmentation reshaped to have few points, onto (e) the same mesh containing more points and finally onto the optimal mesh C_m (b) that we use for our simulations.

navigation system. It can be observed that when using Method 2, a pattern of low values is more spread out than when using Method 1. Comparing the results of the two methods visually, one can also notice some spots on the mesh LV_s representing the values obtained with Method 1. On the contrary, the propagation seems to be more continuous when observing the result of the projection by Method 2. Figure 4.11 illustrates: row (a) the bipolar lat values obtained by intra-cardiac imaging on the RV endocardium, row (b) the projection of these by Method 1 onto RV_s , and row (c) the results of the projection using Method 2. Here the segmented RV_s and CARTO RV_C meshes are very similar. It can be observed that the late values are much better found by Method 2 than Method 1. When trying to compare the obtained projection values with the real values, it is quite complicated to evaluate which method performs best and how to interpret the observations, even in the ideal case when the geometries are very similar (which almost never happens).

Evaluation Method

In order to be able to evaluate the projections methods using more common tools and metrics, we proposed to do an 'inverse' projection. The method to evaluate the projections Methods 1 and 2 consists in using the projection data (i.e., the result of the projection), considering it as the source mesh and projecting its values on the CARTO mesh. Here, the CARTO mesh becomes the target mesh. Once this method is applied, we have for the same original CARTO mesh both the real data and the data resulting from the double projection (i.e., from the CARTO mesh to the segmentation mesh, and from the latter to the CARTO mesh). From these data we propose to calculate: the correlation coefficient, the mean squared error (MSE), the error map, and the standard deviation map.

Results

Figure 4.13 illustrates the results obtained by our method for evaluating the quality of our projection Methods 1 and 2, using the example of a patient's LV CARTO mesh. Figure 4.14 shows another view of the same mesh. In the first column, the invasive measurement of bipolar lat voltages are represented on the LV CARTO mesh (Figure 4.13a). In the second column, we can observe first the projection obtained using the inverse method of projection Method 1, that is projecting from the results of the Method 1 onto LV_s back to the CARTO mesh LV_C . In the same column, below is the standard deviation map obtained from the previous result. In the last row of this column we included the error map resulting from the difference between Figure 4.14b and the map with the original values. The last column of the figure describes the same issue as detailed above for Method 1, but here is exemplified for Method 2. In Figure 4.13d-4.13g (Figure 4.14d-4.14g), small values imply that the projection performs well (small error and small standard deviation), while large values mean that the projection method performs poorly. Looking at the Figure 4.13 and Figure 4.14, it can be noticed that both methods encounter some difficulties in projecting the values inside the mesh. On the contrary, the values are acceptable on the outer surface. This likely due to the fact that there is a large area where part of the outer surface is missing on this mesh. It can also be noted that Method 1 seems to be associated with fewer errors, at least for the projection of the values of this LV mesh. In order to investigate this hypothesis, we compared the results obtained by these two methods onto all the LV and RV meshes we have generated. Based on these error maps and standard deviation maps, we suggest that Method 1 is the most efficient for the projection of Bipolar lat for LV meshes. In addition to these error and standard deviation maps, we calculated the correlation coefficient between the Bipolar lat values of the CARTO meshes and those obtained by the evaluation method for each of the two projection methods. We also calculated the MSE. However, in the context of the voltage values we are trying to project, another measure may be relevant to this study, namely the localisation of the onset of early activation. Therefore, we

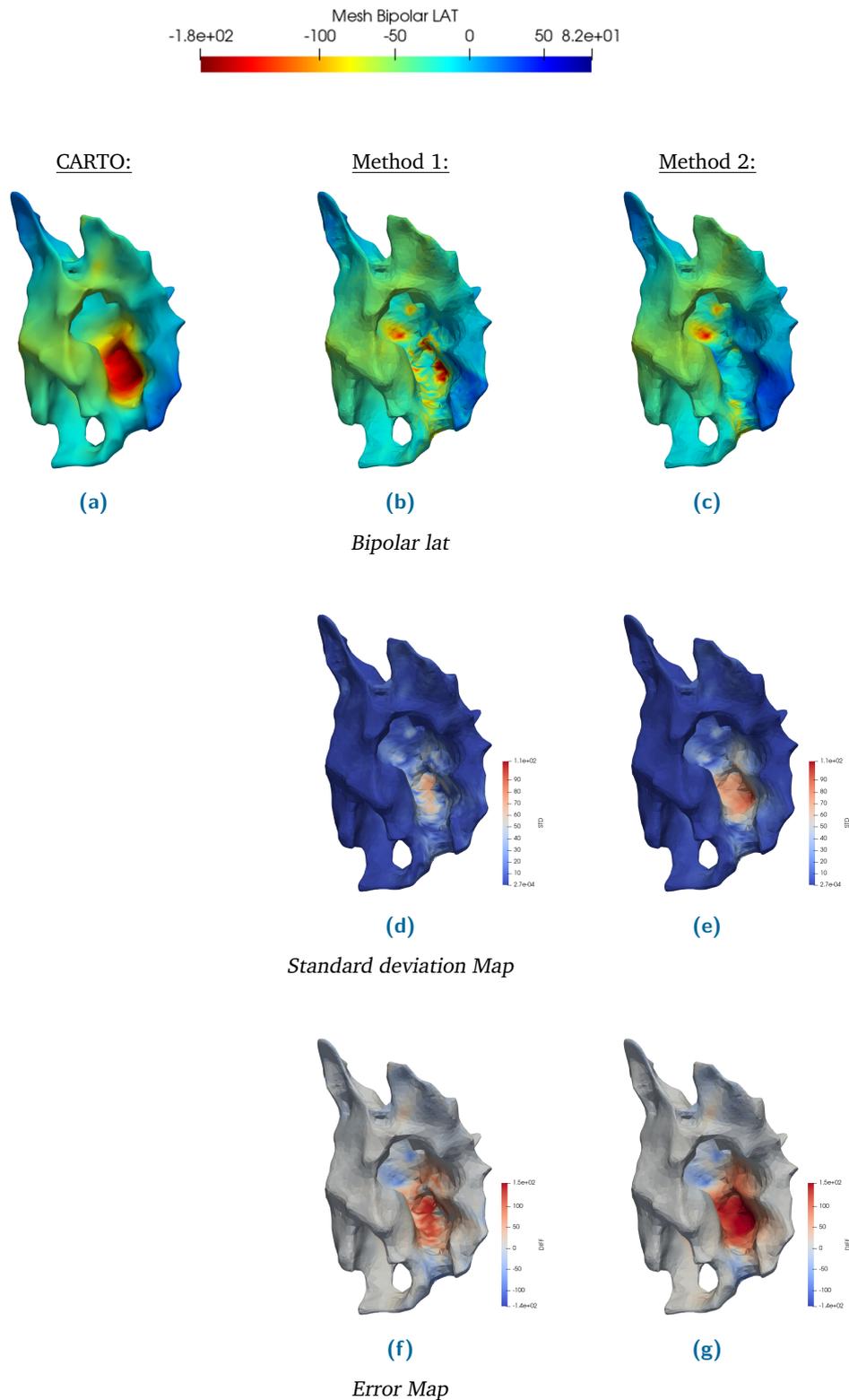


Fig. 4.13: This figure illustrates the evaluation of the methods of projections based on back-projection onto the CARTO mesh: (a) displays the real Bipolar lat obtained from CARTO navigation system, where the second column shows the evaluation of Method 1 and the last column represents those of Method 2; (b) and (c) are the respective projection maps using the inverse Method 1 and Method 2. (d), (e) represent the standard deviation maps calculated between the real CARTO map and respectively (b), (c). In the last row, (f), (g) are the error maps between (b) and (a) and respectively (c) and (a).

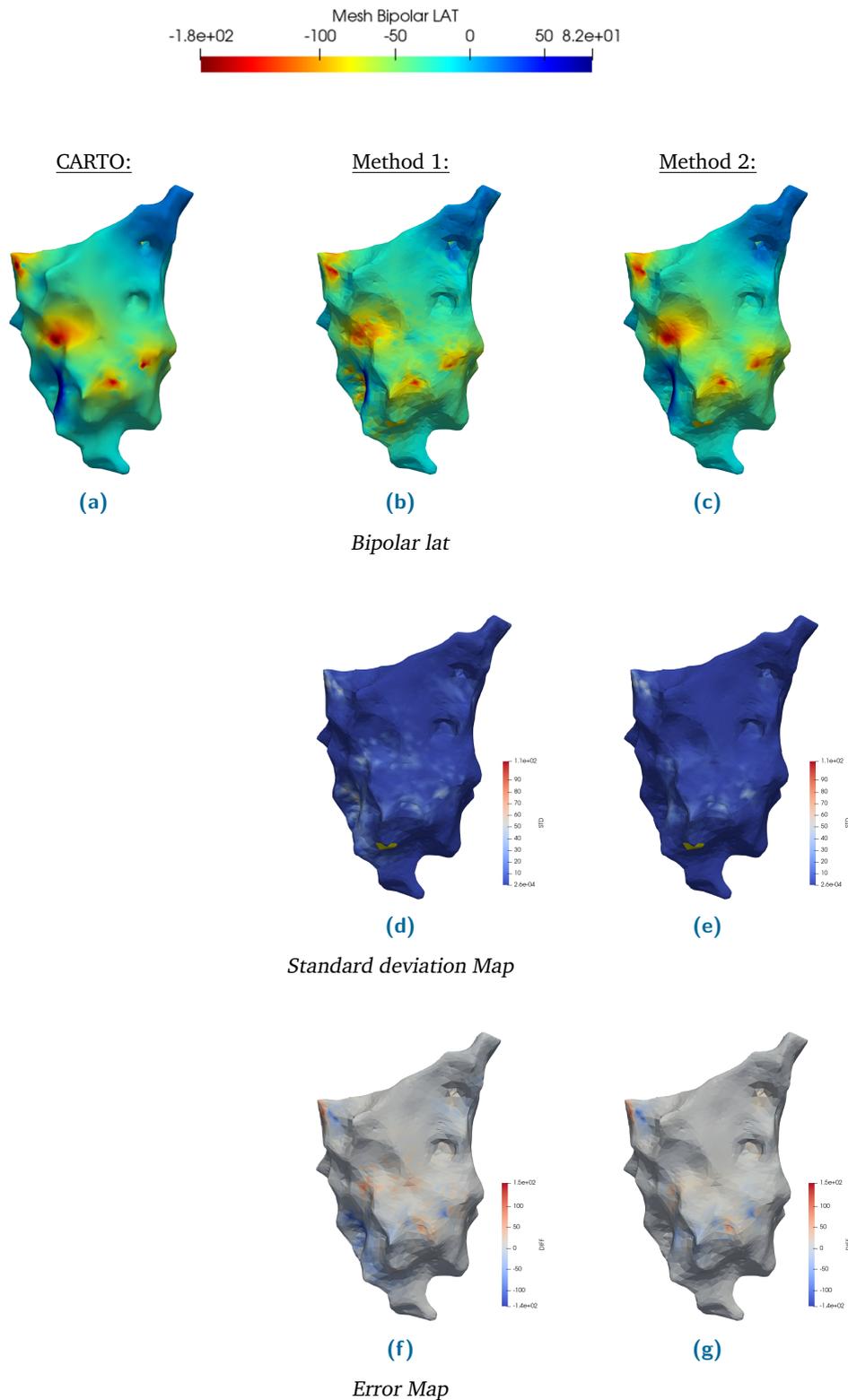


Fig. 4.14: As in Figure 4.13, this figure shows the results obtained to evaluate the proposed projection methods using the same patient case as in the previous figure. Here, the results are presented from a different perspective. (a) is the real bipolar lat obtained from CARTO navigation system. The second column shows the evaluation of Method 1 while the last column represents those of Method 2. (b) and (c) are respectively the projection maps using the inverse Method 1 and Method 2. (d), (e) represents the standard deviation maps calculated between the real CARTO map and respectively (b), (c). In the last row, (f), (g) are the error maps between (b) and (a) and respectively (c) and (a).

have also calculated this measure for each of the maps. To evaluate its corresponding projection error, we calculated the Euclidean distance between the onset localisation on the reference map and that obtained by a double projection.

The results that we interpret and discuss below concern the projections onto LV meshes. To evaluate Method 1, the MSEs obtained are in the range [7.33; 1966] with a mean of 406.0 and a median of 171.1. The correlation coefficient ranges from 0.13 to 0.99 with a mean of 0.77 and a median of 0.83. The calculated mean Euclidean distance is in the range [0.0; 81.5]*mm* and has a mean of 14.7*mm* and a median of 3.4*mm*. The case where the projection performs poorly, corresponds to the LV mesh shown in Figure 4.15. In order to be able to project the values better, it would probably be necessary to pre-process this mesh following the method used for the EPI meshes (see section 4.2.6.1)). The result values given above do not take into account this idea of pre-processing the original mesh. However, in order to understand whether this is necessary and facilitates the projection, we have tested it. Figure 4.15 summarises the results obtained with and without processing. The first row shows the results of the evaluation of the projection Method 1 from the original mesh (Figure 4.15a). The second row shows those obtained from the pre-processed CARTO mesh (Figure 4.15e). Looking at the results in Figure 4.15, starting with the bipolar lat maps, it is easy to see that a whole area of distinct values in the original map (see Figure 4.15a) has not been properly projected when no pre-processing has been done (Figure 4.15b). On the contrary, these values do appear on the map in the projection after processing the original map. Furthermore, the error and standard deviation maps after processing show low values, which means that the projection is rather good. On the opposite, without processing, these maps show higher errors. An improvement of the reference metrics was also observed when processing the mesh: MSE of 280 (previously of 1966), correlation coefficient of 0.87 (before of 0.13) and onset distance of 46 (without pre-processing of 81.5). A significant improvement is observed in favour of the projection using a processed reference mesh.

The same process was applied to calculate the metrics for assessing the quality of the projection by Method 2. The interpretation of the results and our observations allowed us to understand that Method 1 is clearly more efficient than Method 2 when the meshes present many anatomical 'imperfections' (e.g. structures like trabeculation). On the contrary, the difference between the projections is not obvious when LV_C mesh is 'smooth' and is quite similar to LV_S .

The error maps, the standard deviation maps, MSEs, the correlation coefficients as well as the distance of the onset location in turn proved that Method 1 is the most efficient for the projection of LV mesh values. Performing the same research for RV, it was concluded that for these meshes it is preferable to project the values using Method 2. Projection Method 3 has not been evaluated like the two previous ones, since it was developed as a

result of insights and problems that we noticed and analysed when trying to project local activation times (lat) values from *EpiC*.

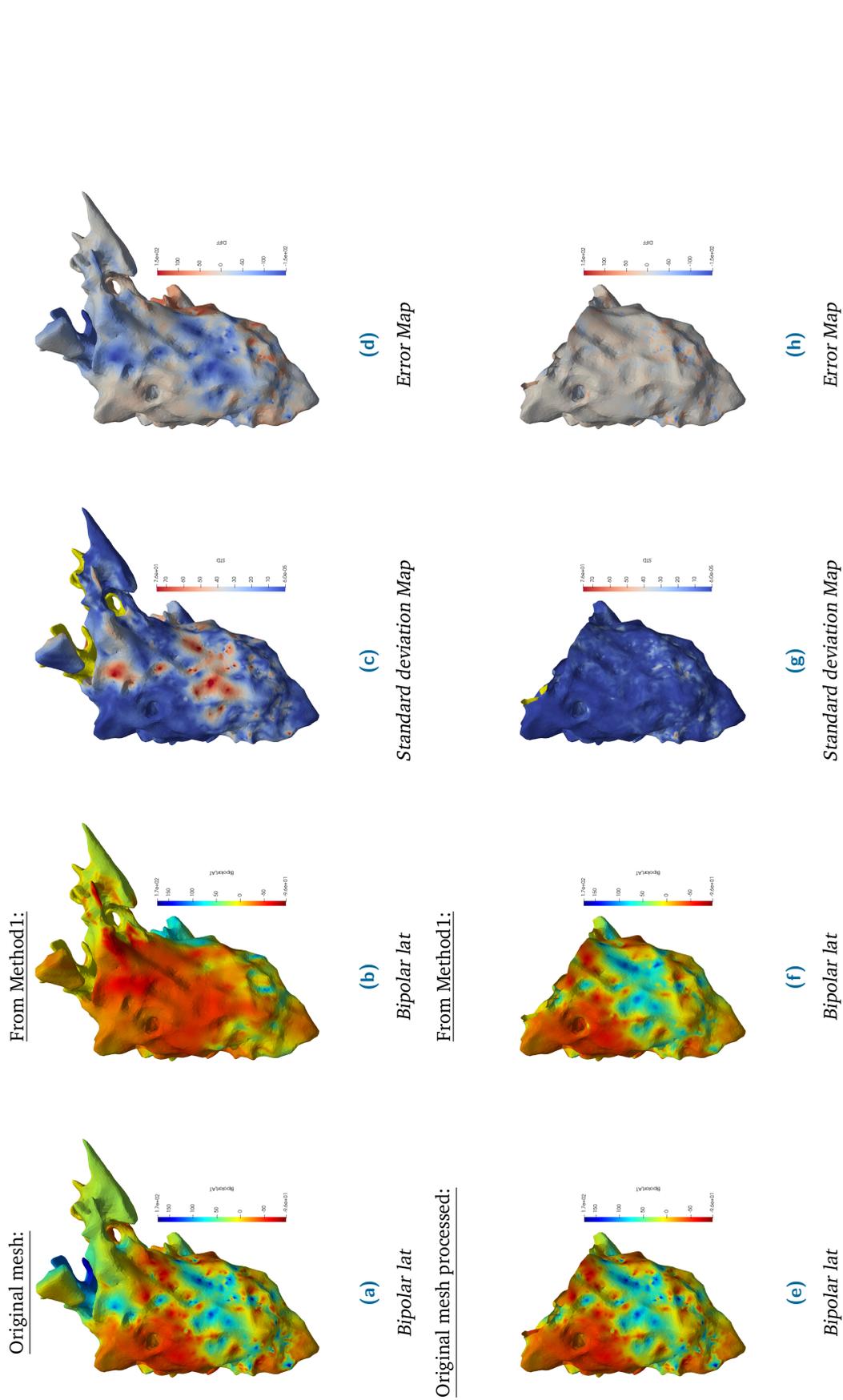


Fig. 4.15: Comparison of the results obtained using the evaluation of Method 1 with processing the reference mesh and without it: (a) reference CARTO mesh with its associated Bipolar lat values, (b) represents the mesh obtained after having performed a processing on the mesh (a). Note that (b) is the result of the double projection using Method 1 and where the values are projected on the mesh (a) and (f) is the result of the double projection using Method 1 and where the values are projected on the mesh (e). (c) and (d) are respectively the standard deviation and error maps calculated from (b) and (a). (g) and (h) are respectively the standard deviation and error maps calculated from (f) and (e).

4.2.6.4 New cardiac mesh with real Bipolar local activation times

The previous proposed strategies allowed us to obtain the optimal cardiac mesh C_m that we will use for our training data but also for our test data. Moreover, we have succeeded in obtaining for each of the available geometries of our CARTO data, its corresponding geometry coming from the segmentation of the CT images onto which the values of the voltages measured invasively were projected. At this stage, we would like to have the data corresponding to the three geometries (LV endo, RV endo and epicardium) integrated into a single mesh, which is the one reconstructed in section 4.2.5. In order to achieve this, there is a first step to be taken, the synchronisation of the maps.

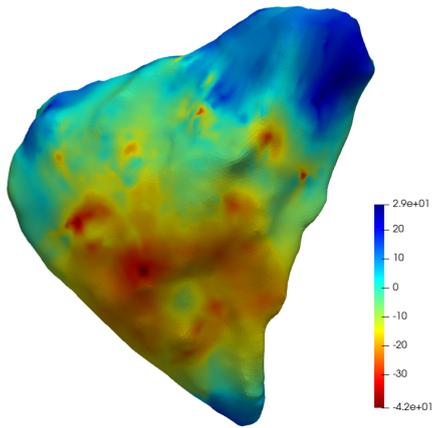
Maps synchronisation

Several approaches to synchronise the bipolar lat maps have been investigated. Indeed, the solution is not unique because the cases encountered are quite diverse. Below we detail the different cases met and the associated method.

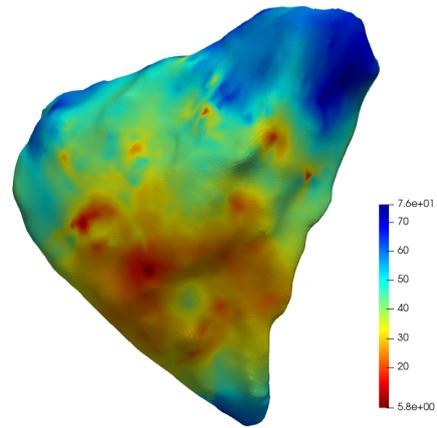
- **Ideal case:** For the same patient, in the ideal case, we include voltage maps for: LV endo, RV endo and epicardium, separately. These three maps of invasive measurements (e.g. the bipolar local activation times we are interested in this work) never start at zero or with positive values, and are shifted in time. Here, the ideal case is considered to be when none of the three maps contain extreme outliers and the values on the epicardium start at a later time than those on the RV and LV. A representation of this is shown in Figure 4.16. It can be seen that the minimum activation time is acquired on the LV mesh, and that this activation time is negative. This was followed by the RV mesh, whereas the last mesh where the acquisitions took place is the epicardium. To synchronise the three activation maps, we need to detect the earliest activation time and add its absolute value to all the values of the three meshes such that the early time after synchronisation corresponds to zero and the other values are coherent with this time shift.
- **Manageable outliers:** As a side note, quite often the bipolar maps obtained from CARTO navigation system contain extreme outliers. This can occur either in the RV map, LV map, or in the epicardial map, or even in more than one map at the same time. In our data, these values are always negative (i.e., no positive outliers were identified). In such cases, the first action to take is to ignore these extreme values and to detect the next early activation time.
- **Challenging and not yet resolved cases:** These are problems/observations for which we have not yet found a plausible explanation. It is currently not possible to propose an adequate method to solve them without the risk of creating new

Before synchronisation:

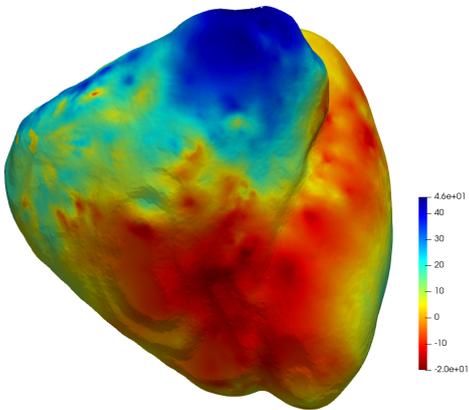
After synchronisation:



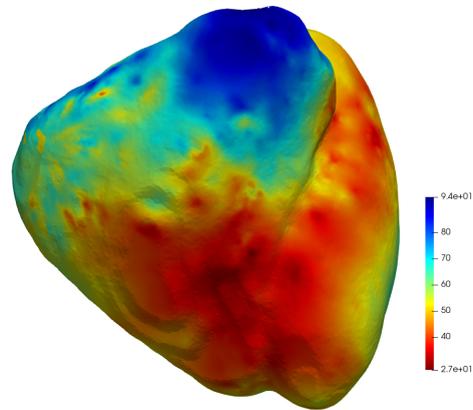
(a) RV



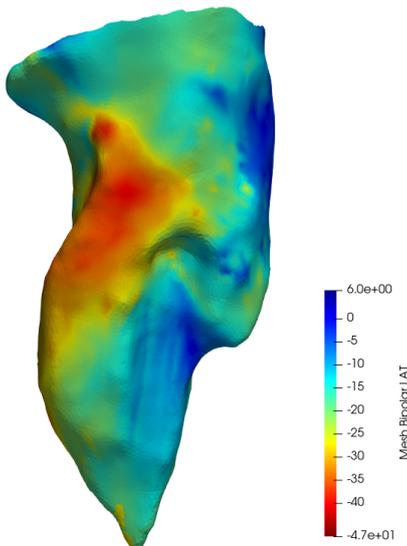
(b) RV



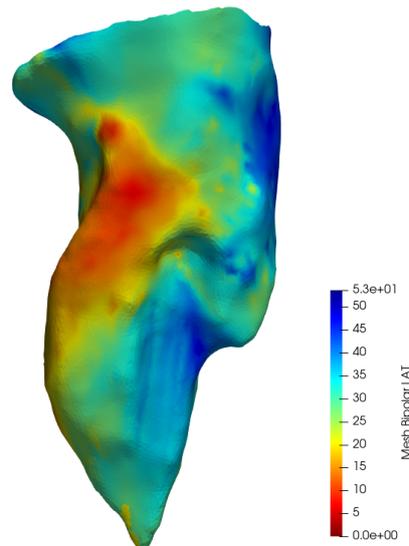
(c) Epicardium



(d) Epicardium



(e) LV



(f) LV

Fig. 4.16: Example of synchronisation of bipolar local activation maps for a so-called ideal case: (a), (c), (e) are the RV endo, LV endo and epicardial meshes, respectively, segmented from CT images, with the associated projected Bipolar lat values, while (b), (d), (f) are the corresponding meshes obtained after the maps synchronisation

meshes containing uninformative values. Information on the different sampling rates could help to solve some of the problems. A list of the different cases that are in progress is proposed below:

- Inconsistency of propagation across the 3 meshes: The wave propagation is from the RV (or LV), followed by the LV (or RV) and finally followed by the epicardium. However, the highest activation time value is not on the epicardium but on the mesh where the activation started.
- Spatial disordering of temporal propagation: the earliest cardiac activation is located in the epicardium. Moreover, the time difference between its value and the shortest RV or LV is greater than 130 ms.
- QRS interval mismatch: the QRS duration of the data does not match that extracted from the BSP data. This duration can be up to 2 times longer.

Optimal mesh with real data

At this stage, we obtained the synchronised maps of the three geometries and we aim to have the values of the bipolar lat voltages on the surface of our optimal meshes (such as the one shown in Figure 4.6d). Previously, we labelled the endocardial LV/RV and the epicardium on our cardiac mesh C_m . Using to these attributes, and by applying a succession of vtk filters, we easily detect the match points of a cardiac geometry to C_m , and associate the bipolar lat value to it.

4.3 Conclusion

In this chapter we have presented the collection of methods that we have applied and that have led to the construction of a database for Electrocardiography. Since our goal was to evaluate our new Deep Learning-based model for solving the inverse problem of ECGI on real clinical data, this data processing step was inevitable.

We hope that by detailing the methods we have implemented and applied in this chapter, it will provide ideas, facilitate and accelerate the implementation of processing, and thus help the clinical data processing community.

Now that our test data set is available, in the next chapter section 5 we use it on an adapted version of our CVAE model presented in the previous section 3.

Evaluation of Deep Learning ECGI formulation with Clinical Data

Contents

5.1	Introduction	83
5.2	Materials and Methods	85
5.2.1	Clinical Data	85
5.2.2	Data Optimisation	87
5.2.3	Data sets Construction	88
5.2.4	Deep generative model for ECGI	91
5.2.5	Computational Details	93
5.3	Results	94
5.4	Discussion	99
5.5	Conclusion	101

In section 3, we presented a new approach based on Deep Learning to reformulate the entire problem of ECGI as a β -Conditional Variational AutoEncoder based on convolutional neural networks. The ability of this approach to estimate maps of cardiac activation throughout the myocardium in a volumetric manner was demonstrated from non-invasive data corresponding to BSPs, which we simulated by modelling the electrode vest on the torso with a 10×10 electrode Cartesian grid, and from myocardial masks obtained by segmenting CT images of patients. In order to evaluate this model, we have developed an approach to solve the inverse problem of the ECGI, and then proceeded to processing all clinical data provided to us by the CHU Bordeaux. This work has been detailed in section 4. However, considering that there is a significant difference between real and simulated data, we have made several modifications to our previous formulation [Bacoyannis, 2021] in order to learn more extensive information. Methodological improvements over the preliminary method mainly include conditioning data and data generation. These changes are presented in the method section 5.2. In the following section 5.3, we will present the results we have obtained with our new formulation on synthetic data and by evaluating it on clinical data. Finally, we will discuss our preliminary results and outline limitations and future work in conclusion.

Abstract:

Aims Providing clinicians with non-invasive cardiac activation maps of superior quality would allow them to manage more patients, make earlier diagnosis, guide interventional therapy, and perhaps also predict response to therapies such as cardiac resynchronisation and ablation. Electrocardiography is a non-invasive clinical modality for mapping the electrical activity of the heart using surface electrodes and body surface potentials (BSPs). However, the method for solving the associated inverse problem (known as ill-posed) still needs to be improved. Previously, we proposed a deep learning based method that proved to be effective and convincing in solving the inverse problem for purely synthetic data by leveraging multi-modalities information and prior knowledge from previous cases. Here, we take one step further and propose a probabilistic model that allows us to generate activation maps from real BSP data.

Methods To construct both training and evaluation datasets, we proposed a data preparation pipeline. First, the clinical data from 11 patients were processed, then a geometrical modelling of the torso and the heart was proposed, on which: 1) a fast personalised cardiac simulation pipeline from the eikonal model is applied in order to build the training dataset; and 2) the ground truth data (i.e., activation maps from CARTO data and BSPs from ECGI data) were transferred to create the evaluation dataset. In addition, processing of the BSPs to fit the simulated data to the real data has been performed. The DL-based model for ECGI inverse problem is a Conditional Variational Autoencoder (CVAE) based on convolutional neural networks. The main idea of our CVAE is that it learns the latent distribution of the given data during the training, so that new meaningful samples (cardiac activation maps) can be generated from it combined with the conditioning data (e.g. BSPs and propagation conduction velocity maps).

Results Our model was trained on five different cardiac anatomies (resulting in 4500 activation maps). At each epoch, the validation dataset is of 200, randomly selected from a set of 1800 maps. This novel model was evaluated by testing it on two additional patients. The comparison of our generated activation maps with the real data shows a good identification of the cardiac activation pattern. The calculation of the distances of 7 onsets (mean:11.9 mm), indicates a good predictive power on the location of the latter (usually between 10 and 20mm in the literature).

Conclusion We showed that by probabilistically learning spatio-temporal correlations and naturally including anatomical information, from non-invasive multi-modal data, simulated with a fast and simplified electrophysiological model, we are able to

generate good predictions of cardiac activation from non-invasive clinical data (BSP and cardiac geometry).

Significance Being able to generate in 1 second an activation maps from a combination of real non-invasive data such geometric data and torso surface signals, with good prediction of activation patterns and onsets locations, is a step forward in the field of ECGI.

Keywords: Electrocardiography, Inverse Problem, Deep learning, Computational Modelling, Generative Model, Clinical Evaluation

5.1 Introduction

Activation maps are a necessary source of information for clinicians to diagnose cardiac arrhythmia or to guide therapeutic interventions. Currently, the conventional routine method for obtaining this type of maps is based on an invasive procedure, namely the electro-anatomical mapping (EAM). This procedure consists on introducing a catheter into the patient's heart chamber and acquiring point to point intracardiac electrocardiograms by moving the catheter tip over the heart endocardial and/or epicardial surfaces. After the invasive mapping is performed, the voltage potentials measured at each of these points are interpolated to obtain a map of local activation times. An alternative non-invasive method is Electrocardiographic Imaging (ECGI). ECGI estimates an activation map from three-dimensional anatomical data of the torso and heart geometry obtained from CT imaging combined with voltage potentials that are measured on the torso surface by a multi-electrode vest.

The mathematical process for reconstructing these activation maps is the inverse problem of ECGI. Most of the formulations of this inverse problem are known to be ill-posed in Hadamard's sense. Typically, at the expense of accuracy loss, most approaches solve the ECGI problem by employing regularisation techniques for the stability of the solutions. The classical methods of solving it have not yet allowed ECGI to lessen the EAM usage. More details about the classical methods can be found in [Karoui, 2018; Cluitmans, 2018].

In recent years, much attention has been directed to the use of machine learning models to solve the inverse problem of ECGI. To our knowledge, the first proposal in the literature came in 2010 [Jiang, 2010], where support vector regression (SVR) [Vapnik, 1999] was used to reconstruct transmembrane potential (TMP) on the epicardial surface. The reconstructions produced by this method are more accurate than those produced by the Tikhonov regularisation. In [Giffard-Roisin, 2018], authors suggest a data-driven regular-

isation approach where a reference torso anatomy onto which the BSPs are projected before using an relevance vector regression method (to solve the inverse problem) and the reconstructions obtained are cardiac activation maps. A Variational AutoEncoder was used in [Ghimire, 2018] to reconstruct the TMP, while in [Toloubidokhti, 2021] a VAE-based method aimed at reducing the errors in the forward model. In recent research studies, machine learning-based models have been also investigated to localise ventricular activation, tachycardia origin and premature ventricular contraction (PVC) origin from 12-lead ECG [Missel, 2020; Arrieula, 2021].

A suitable method to properly solve ECGI inverse problem is still topical and at the forefront of research. In this context, in [Bacoyannis, 2021], we proposed a novel method based on a β -Conditional Variational AutoEncoder able to solve the inverse problem of ECGI. This generative probabilistic model learns geometrical and spatio-temporal information, and enables the generation of 3D volumetric activation maps for a new cardiac geometry from non-invasive BSP and cardiac imaging data. We should also highlight that the correlation between the imaging substrate and the signals is also learned, allowing us to integrate any 3D image information into the ECGI-based model. This is rather difficult to achieve in the classical formulation; however, such previous work can be considered as a proof of concept, in the sense that it has been validated on synthetic data.

In continuity with our previous work, here we conduct an evaluation study on real clinical data. The specific aim of this study is to evaluate the performance of estimating activation maps from real non-invasive ECG signals on the torso. However, to enable this, it is first necessary to render the evaluation possible. Evaluating and comparing multi-modalities data of different geometries that was imaged at different times remain challenging tasks. When it comes to developing a deep learning model, preparing the data set is the first and not least step in the process. In our case, these data are on one hand real data that are complex, come from different modalities and on the other hand there are data that we have simulated. The objective of data preparation is to format these data and to create data sets that will be ready to be used by our learning model. This objective gives rise to a subsequent step of clinical data processing, which will be described in section 5.2 of this chapter. In the same section, we describe the new features adopted for the generation of our simulated database. We then present our new formulation of the whole ECGI problem, which is derived from the previous one tested on synthetic data, which is now more adapted to real data. In section 5.3, we present the preliminary results we have obtained. To conclude, we will discuss the results and present our ideas for possible future improvements.

The different contributions of this manuscript are:

- A processing framework for ECGI clinical data. This provides a clinical database of 11 patients. It consists of geometric data of the heart and torso, body surface potentials and bipolar local activation times on the segmented cardiac mesh.
- A novel simplified personalised torso potential simulation model, allowing us not only to accelerate the speed of simulation of BSPs, but also to reduce the computational time of the training of the DL-based model.
- Adding the propagation conduction velocity map to the simulation pipeline, even more specifying constrain on it and use it as conditioning data in our model, allows to better/more realistically reproduce the cardiac activation pattern.
- Applying processing to our simulated BSP so that they are more consistent with real BSP, while keeping the information content, would allow our model to learn correlations that link activation maps to signals closer to reality.
- A simulated database composed of 9900 activation maps and their corresponding BSPs. Three scenarios for simulating cardiac activation maps have been considered: one onset on the right ventricle, one onset on the left ventricle, three total onsets (1 on LV plus 2 on RV). All of the initialisation points were sampled randomly (from a discrete uniform distribution) in accordance with the different scenarios.
- On learning from statistical correlation between ECGI simulated data, a generative β conditional variational autoencoder was proposed, allowing us to generate new maps of the cardiac activation by providing it only a new propagation conduction velocity map and the noninvasive real BSP.
- As a result of its architecture based on the CNN, our model is able to learn spatio-temporal correlations as well as the influence of cardiac shape/structure.
- A probabilistic model trained on 4 500 data, ready to be used as soon as new clinical data become available, to generate its corresponding activation maps.

5.2 Materials and Methods

5.2.1 Clinical Data

In this study, we handled clinical datasets from 11 patients provided by the CHU Bordeaux. Each patient dataset consists of three separately collected components: the Computed Tomography imaging data (to derive the myocardial geometry C_m), the inva-

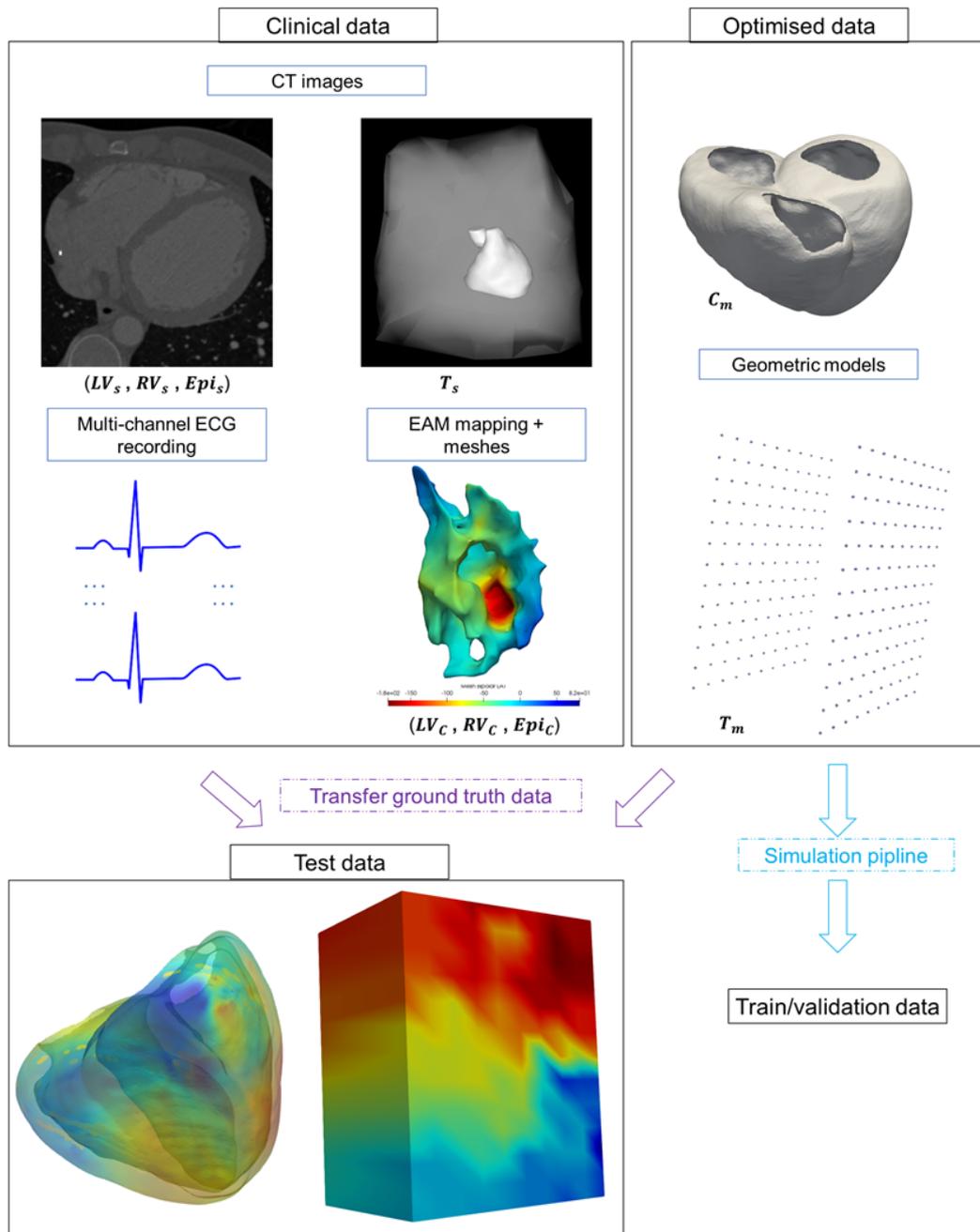


Fig. 5.1: Representation of the main steps in obtaining the required training/validation/test data sets for our Deep learning based model for ECGI. From the clinical data, geometric models (i.e., C_m and T_m are developed and will be used for two purposes: 1) by applying our simulation pipeline on the proposed torso and cardiac models, cardiac activation maps and their corresponding BSPs are obtained, which will constitute the training/validation data sets; 2) test data are obtained by projecting the clinically acquired data (BSPs from ECGI data and bipolar lat from CARTO data) onto the geometric models that we have proposed.

sive mapping data from EAM (resulting in bipolar local activation times and voltages on

LV_C, RV_C, Epi_C and for which we have sinus rhythm maps), the ECGI data involving CT images (allowing the torso geometry T_s and the ECG-leads localisation) and BSP data from the CardioInsight technology.

The clinical data processing includes: acquisition and segmentation of CT images, measurement and analysis of non-invasive ECG signals on the torso, as well as of invasive catheter-based mapping of activation times and unipolar/bipolar voltages via a conventional electro-anatomical system.

5.2.2 Data Optimisation

Anatomical Cardiac model

We generate 3D cardiac meshes from segmented CT images that were obtained using an automated segmentation method based on deep learning [Cedilnik, 2019]. The cardiac model provides anatomical information about the heart, and this is also the support used to simulate our synthetic activation maps and the model on which invasive EAM data will be projected onto. In order to use this cardiac mesh to generate our database of activation maps, we first proceed to the elimination of unnecessary anatomical structures such as papillary muscles and valves, and finally we label the endocardium, epicardium as well as the right/left ventricles. After a uniform remeshing using the clustering approach from ACVD library [Valette, 2008], each optimised cardiac mesh C_m is a polygonal mesh of 50K vertices. Because the meshes reconstructed after the invasive procedure correspond to different times than the meshes generated from the CT images, they do not belong to the same space and are not oriented along the same axes. Therefore, we proceed to a manual registration of meshes using landmarks via the MUSIC[®] platform. The reference frame we selected here is the one corresponding to the generated mesh C_m .

Torso model

The torso is modelled for two reasons: firstly to have anatomical information and secondly to have a model on which to simulate the BSPs. The thoracic meshes obtained were segmented by the method based on the seed point. Further to the torso model proposed in [Bacoyannis, 2021], we have made new improvements. Here, we model the torso by two grids of electrodes representing the anterior and posterior surfaces of the torso. Notably, these are 10×13 electrode cartesian grids. Thus, our model T_m comes as close as possible to the 252 electrodes of the CardioInsight technology vest. The torso model is further customised for each patient. More specifically, the distance between two consecutive electrodes to take into account the morphology of the patient's body as well as the location in space so that the torso model T_m and the real torso T_s , are represented

in a unified reference system. Figure 4.4 is a representation of T_s and T_m in this basis, where the vectors $(\vec{0x}, \vec{0y}, \vec{0z})$ form an orthonormal Cartesian coordinate system with direct orientation. The positions of the actual electrodes are extracted from T_s , from which the minimum and maximum positions are calculated along the three axes x, y and z. The minimum and maximum positions along the y-axis allow us to fix the position of our two-electrode grids. From the x-positions, we determine the inter-electrode distance in the direction representing the width of the torso, while from the z-axis locations, we can modulate the height of the modelled torso T_m .

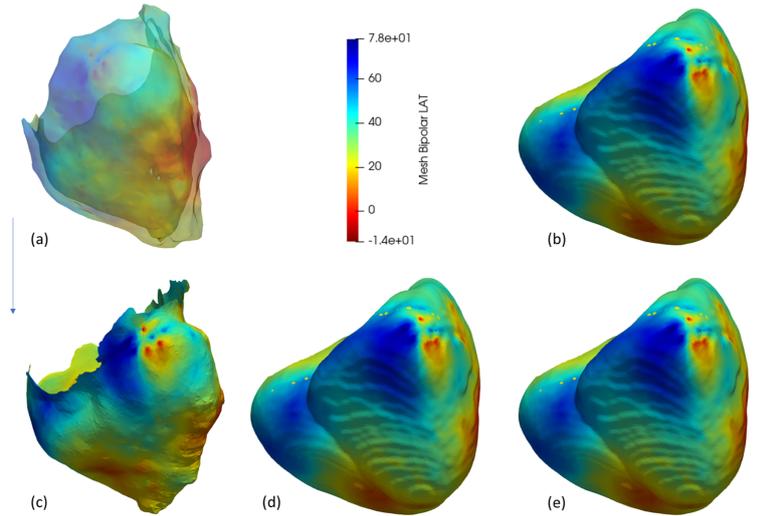


Fig. 5.2: Detailed representation of the process used for the projection of (a) local activation times from Epi_C , the CARTO mesh of the epicardium to (b) the volume mesh used for our simulations. The projection from (a) to (b) is done via intermediate steps: first the LAT points are projected onto (c) the surface extracted from the reference mesh, then onto (d) the mesh of the segmentation reshaped to have few points, and onto (e) the same mesh containing more points and finally onto the optimal mesh C_m (b) that we use for our simulations.

5.2.3 Data sets Construction

5.2.3.1 Test data set: transfer from clinics to anatomical models

Cardiac activation maps

For evaluation purposes, we need the activation times of ground truth CARTO data projected onto the optimised meshes C_m . Once the meshes obtained invasively from CARTO navigation system are registered, their bipolar lat (i.e., local activation times) values are projected onto the optimised meshes. From the invasive data obtained by CARTO, the surface meshes of the RV, LV and the epicardium are reconstructed independently (i.e., we do not have a single volume mesh of the heart with all associated measured values). In order to be able to project the values of the latter (i.e., LV_C, RV_C, Epi_C onto our generated optimised mesh C_m , we use the meshes generated after segmentation of

the LV (LV_s), RV (RV_s) and epicardium (Epi_s), independently. Depending on whether we want to project the bipolar lat values of the either LV_C , RV_C , or Epi_C , the projection method will be different. However, all three methods are derived from each other and are based on the principle presented in [Cedilnik, 2018]. In all three approaches, we let an ellipsoid fit the mesh vertices, calculate the centre of the frame and redefine all points coordinates into spherical coordinates. To project the values from invasive source mesh onto the corresponding target mesh, we use a linear interpolation method. For RV mesh, the projection is performed after having first computed the normals. In this case, the values of the internal surface of the mesh are projected onto RV_s . Finally, regarding the epicardium, the same method as for the RV is applied, but intermediate steps (as shown in Figure 5.2) are taken so that the projection is not too abrupt. We tested different projection methods in order to develop the latter. However, as our research was more result-oriented, we did not carry out a benchmarking of the different projection methods. It would be interesting to propose it as in [Soto-Iglesias, 2016].

Body surface potentials

In order to have the real BSPs on our personalised torso model T_m obtained as explained in section 5.2.2, we calculate the Euclidean distances between the electrode locations of the real mesh T_s and our model T_m . We then infer the matching electrode for each electrode belonging to the grids. Thus, each electrode in the model is associated with the corresponding voltage (potential) value of the matching electrode. However, as our Cartesian grids have 8 more electrodes than the CardioInsight vest, it is not a 1 to 1 match.

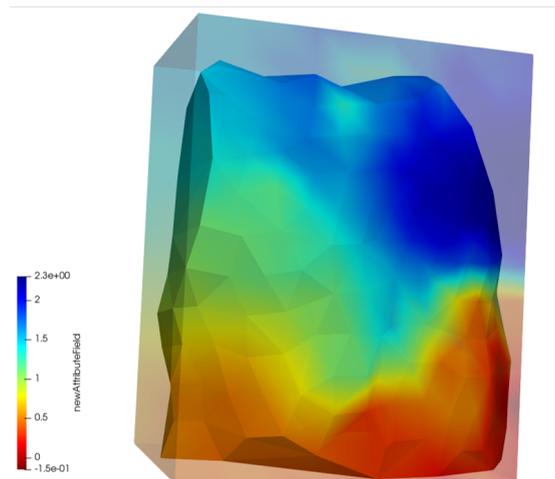


Fig. 5.3: Illustration of the Body Surface Potentials on the original torso mesh T_s and on the corresponding mesh of the torso model T_m .

5.2.3.2 Data Generation: training/validation data sets

We used a fast personalised cardiac simulation pipeline [Desrues, 2021] to produce our activation maps and the corresponding BSP. Activation maps are generated using the Eikonal wave front propagation model, as in our previous work. However, here we simulate data considering multiple onsets. For each optimised cardiac mesh C_m resulting from section 5.2.2, we simulate: 200 activation maps considering single pacing sites on RV; and, 200 activation maps considering 2 onsets localised on LV. Additionally, 500 activation maps are simulated considering 3 onsets (2 being on the RV and 1 belonging to the LV). In line with [Giffard-Roisin, 2017], we used the anisotropic Mitchell-Schaeffer cardiac model with the dipole formulation to generate BSP on T_m .

Adaptation of the simulation to make it more realistic

- **Cardiac activation map:** Furthermore, compared to the previous pipeline used to simulate activation maps, a thin endocardial layer ($< 2mm$) is now assigned to a higher isotropic propagation velocity. This allows to better reflect the fast propagation of the action potential generated by the Purkinje network. Myocardial propagation velocity is set to $0.5m.s^{-1}$ while endocardial propagation velocity is set to $1.25m.s^{-1}$ (including the junction between Ventricles and Artrials). An example of propagation conduction velocity map is represented in Figure 5.4.
- **BSP:** As shown in Figure 5.4(c), the clinically measured potentials differ from the simulations notably in terms of their amplitude and area under the curve. One main difference we observed between our simulated signals and the real ones, is an amplitude shift. While our simulated signals appear to be shifted to zero, the real ones are not. In order to adjust our signals, we add a baseline wandering noise to those values. However, this is not systematic. This noise is added to the signals during the training phase of the model. The training is done with the signal without noise and with noise. Depending on whether the loss reconstruction is better with or without the noise, we choose to keep the modified signal or not. Furthermore, as the localisation of the reference electrode is not communicated, we have used the Wilson's central terminal (WCT) instead. The WCT allows a reference point to be constructed artificially. It is equal to the average of the three electrodes connected to the right arm, left arm and left leg. Once the average potential is obtained, we subtract its value from the potentials of each of the electrodes. Figure 5.4(c) shows the real signals, the raw simulated signals and the simulated signals after adaptation using the method described above.

In total, we built a synthetic database of 11 patients containing 9900 activation maps and as many BSP values.

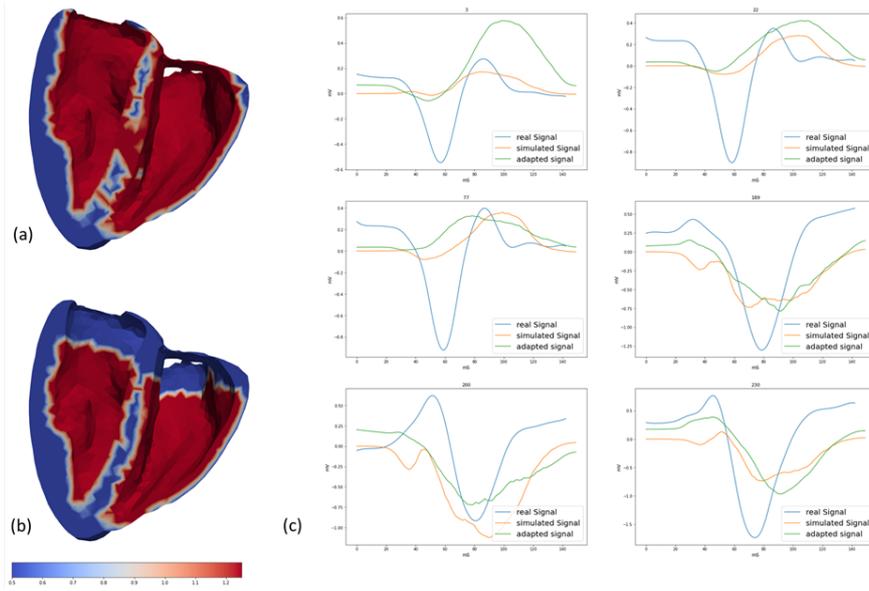


Fig. 5.4: (a) Example of propagation conduction velocity map where the myocardial propagation velocity is set to $0.5m.s^{-1}$ while endocardial pro is set to $1.25m.s^{-1}$, (b) the propagation conduction velocity map after including the junction between Ventricles and Atria), (c) Illustration of the signals at 6 different electrodes. The real signals are in blue, the raw simulated BSPs are in orange, and the simulated signals that we have adapted to the real signals (addition of noise + WCT) are shown in green.

5.2.4 Deep generative model for ECGI

A Conditional Variational Autoencoder (CVAE) [Kingma, 2014c] is an extension of VAE [Kingma, 2014b] which is an autoencoder with a distribution over the latent space, what makes it a generative model. In the case of CVAE, the latent variables z and the data x are both conditioned on additional random variables c . The c variables condition not only the input data of the encoder, but also the decoder. The resulting variational probability becomes: $Q_{\Phi}(z|x, c)$, while the generative distribution of the decoder is: $P_{\Theta}(z|x, c)$. We assume the prior distribution $P(z)$ to be a centred isotropic multivariate Gaussian $N(z : 0, I)$. The CVAE is trained to minimise the loss function (Equation 5.1) which comprises two terms: a reconstruction loss and the Kullack-Leibler (KL) divergence which is a measure between the variational probability and the true posterior distribution. By minimising the loss function, the evidence lower bound on the probability of generating real data samples is maximised.

$$L_{CVAE}(\Theta, \Phi) = -\log P_{\Theta}(x|z, c) + D_{KL}(Q_{\Phi}(z|x, c) \parallel P(z)) \quad (5.1)$$

The KL divergence (D_{KL}) is:

$$D_{KL}(Q_{\Phi}(z|x, c) \parallel P(z)) = \sum Q_{\Phi}(z|x, c) \log \frac{Q_{\Phi}(z|x, c)}{P(z)}$$

Commonly, $Q_{\Phi}(z|x, c)$ is Gaussian with a diagonal covariance matrix: $z \sim Q_{\Phi}(z|x, c) = N(z; \mu, \sigma)$ where: $z = \mu + \sigma \odot \epsilon$, \odot denotes the Hadamard product and $\epsilon \sim N(0, I)$. The D_{KL} term can be computed analytically (see appendix B in [Kingma, 2014b]).

According to [Higgins, 2017], we introduce the penalty coefficient β to the KL term. It allows to adjust the relative weight between the reconstruction and the regularisation terms. Introducing β into the CVAE intends to discover disentangled latent factors.

In [Bacoyannis, 2021], we modified the loss reconstruction to take into account our data and the problem we aim to solve. A detailed explanation is available in section 3. The proposed loss function to minimise when training our model can be written as:

$$L_{ECGI} = L_{\beta\text{-CVAE}} + L_z + L_{\bar{z}} \quad (5.2)$$

where: $L_{\beta\text{-CVAE}}$ is the KL divergence loss, L_z and $L_{\bar{z}}$ are loss functions based on the two modified versions of Root Mean Squared Error (RMSE) we developed. L_z is the reconstruction loss between the encoded data and the ground truth, whereas $L_{\bar{z}}$ is the reconstruction loss from the multivariate Gaussian.

The architecture of our improved Deep Learning model, proposed in this thesis, is described in Figure 5.5.

Our previous Deep Learning model dedicated to solve the inverse problem in 3 dimensions has been revised as described in following section.

- The conditioning data of our model are actually: the BSP of the new torso model and the propagation conduction velocity map of the heart.
- The original dimensions of the input data are: [120, 120, 120] for the activation/propagation conduction velocity maps, [260, 200] for the BSP.
- A pre-processing step of the signals data (called 'signal corrector') is required and applied to the original signals such that the output BSP-s are of the same size as the other inputs data. This signal corrector process is included in the β -CVAE model and is based on three steps: 1) first, we reshape the initial signals of dimension equal to [260, *signal_length*] into [20, 13, *signal_length*]. Currently, we cut the input signal length to 200ms; 2) then the signal is pad to [20, 15, 200]; and 3) finally, we modify the resolution at [120, 120, 120] by applying three consecutive 2D transpose convolutional layers.
- The conditioning data of our β -CVAE are obtained by a concatenation of the signal (pre-processed as described above) and the mask of cardiac propagation velocity.

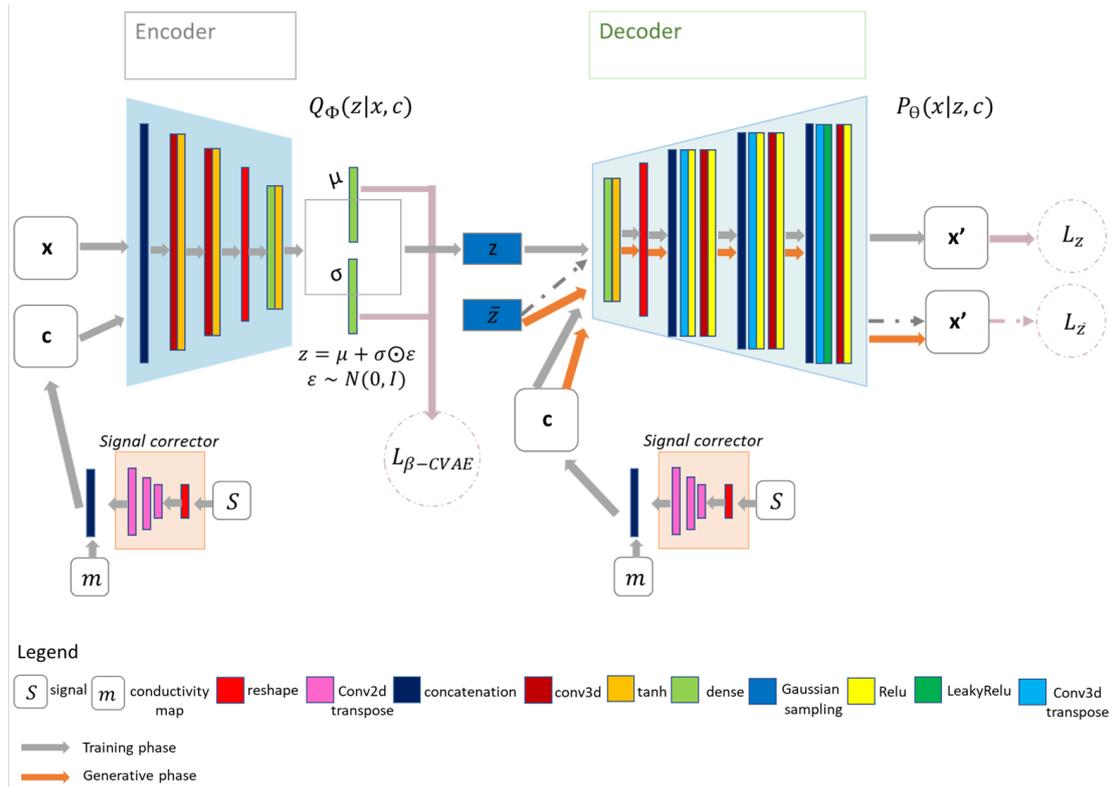


Fig. 5.5: β -CVAE model architecture. *Training:* The input signal is reshaped before changing its resolution and becoming of the same dimension as the propagation conduction velocity map, by applying three consecutive 2D transpose convolutional layers. Then the propagation conduction velocity map and the resulting signal from the 'signal corrector' process are concatenated to form the condition input. Total input data is condition input concatenated with activation map and encoded in z . Activation maps are decoded from z with conditioning by sub-sampled versions of the BSP mapping signals and of the speed mask. *Prediction:* decoder is used to generate myocardial electrical activation maps from a propagation conduction velocity map corresponding to a new cardiac segmentation, and its corresponding mapped BSP signals.

5.2.5 Computational Details

Our data-driven generative model was implemented using Keras API of Tensorflow 2.0¹. Using Adam [Kingma, 2014a] optimiser with a learning rate of 10^{-4} , we trained the model for 240 epochs with a batch size of one.

The model was trained on five different patient geometries, including a total of 4 500 activation maps and their corresponding BSP. At each epoch, the validation dataset is of 200, randomly selected from a set of 1800 maps.

The latent variable size was set to 25. The model was optimised using the total loss described in Eq. 5.2, without any additional weight applied to the loss terms. The best value for β , the penalisation term in the loss function, was 10. The total computation

¹https://www.tensorflow.org/api_docs/python/tf/keras

time for training was 64 hours on a NVidia Tesla V100 (32GB of RAM). The trained model was tested on 200 activation maps and their corresponding BSPs from two unseen geometries. To evaluate the performance of the model on unseen new patients data, we calculate: the mean absolute error between the ground truth activation map and the mean activation map over 10 predictions, the Euclidian distance between 7 onsets location and their temporal distance.

To adapt the signal shape, we used three 3×3 2D convolutional layers with the number of strides set to $[1, 2]$, $[2, 2]$ and $[3, 2]$ and the number of output features to 200, 200 and 120, respectively. The encoder of our network consists of two successive stride 2 convolutional layers, one flatten layer followed by one dense layer. Each layer has a $3 \times 3 \times 3$ kernel and all the activation functions are \tanh . The bottleneck layers (μ, σ, z) are fully-connected. The decoder has a fully-connected layer which is then reshaped into a 3D matrix (excluding the batch and channel dimension) before applying a succession of transpose and normal convolutional layers.

To adapt the simulated signals to the real BSPs, we added the baseline wandering noise to the simulated signals when training the CVAE model. To add the noise, we add the randomly generated low frequency ($0 - 0.5Hz$) sinusoidal noise for each electrode.

5.3 Results

To evaluate our method we used a dataset with 100 activation maps and their corresponding BSPs per patient. We tested the model on the anatomies of two additional patients. We performed an evaluation on synthetic data and then on real data. Because of the probabilistic properties of our generative model, for each case it is possible to generate multiple activation maps (i.e., distinct predictions), by sampling the latent space. For each of these 100 different BSPs, we have chosen to generate, 10 predictions (10 activation cards). Next, from these ten predictions, we calculated the mean map and the standard deviation map which will be used to compare our predictions with the ground truth activation.

Cardiac activation from synthetic BSP:

In this section, the activation maps were generated from simulated BSP data. Figure 5.6 show the results obtained for #Patient10, whose data were never seen during the training of our model. In this figure: (a) shows the ground truth map in two different perspectives; (b) represents the activation map corresponding to the mean calculated over 10 predictions; and (c) corresponds to the standard deviation map calculated also over these 10 predictions. As we can observe in Figure 5.6, our probabilistic model allows

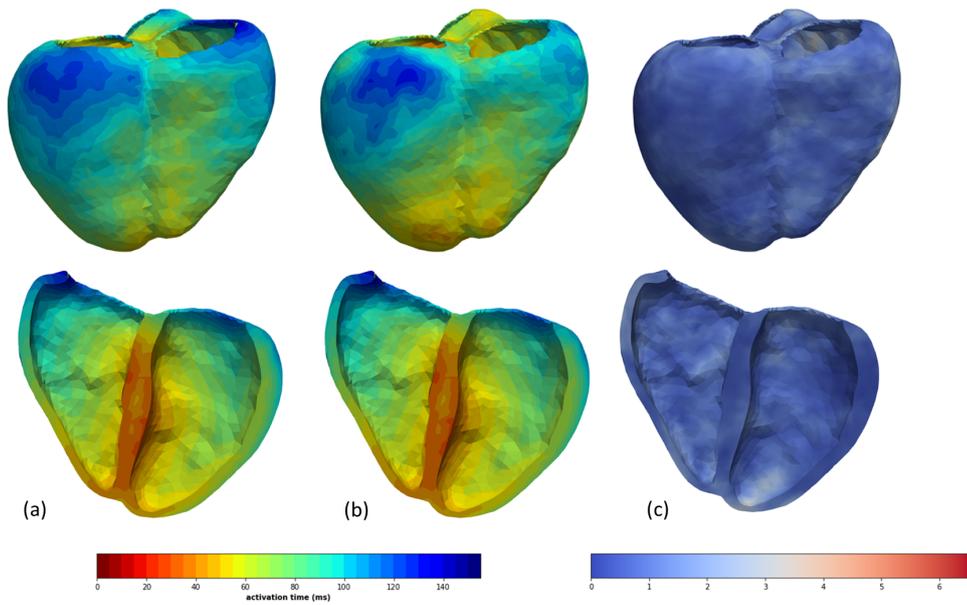


Fig. 5.6: Results obtained by testing the model on a new geometry (#Patient10) with synthetic EP data. (a) ground truth activation map; (b) mean activation map over 10 predictions, generated by our model; and, (c) standard deviation map calculated over 10 predictions.

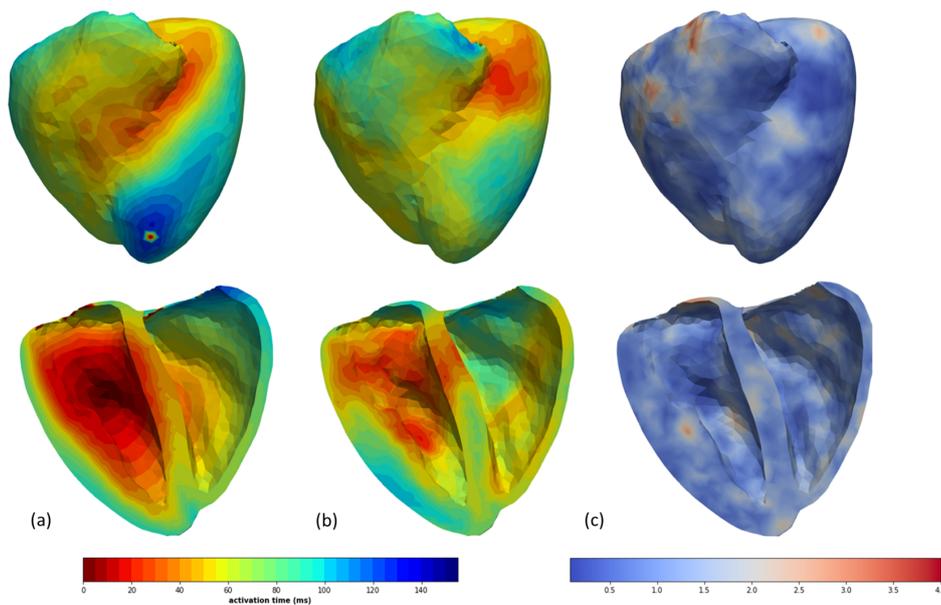


Fig. 5.7: Results obtained by testing the model on a new geometry (#Patient12) with synthetic EP data. (a) Ground truth activation map, (b) mean activation map over 10 predictions, generated by our model, (c) standard deviation map calculated over 10 prediction.

us to generate new volumetric activation maps with a good accuracy when generated using simulated BSP as conditioning data. The mean absolute error (mae) corresponding to the case presented here is 14.4ms with standard deviation (std) of 3.81 ms. Note that it is assumed that the mae and std have been calculated only in the region of interest.

Figure 5.7 is similar to the previous one but concerns the activation generated for #Patient12. The mean absolute error over the 100 tests is 18.6 ms with 2.34 ms standard deviation.

Cardiac activation from real BSP:

Ground truth activation maps and activation maps generated from clinical BSP were compared in terms of mean absolute error and onset error (physical distance and temporal delay). Our model generates volumetric activation maps. In order to compare our predictions with clinical data, which are surface-based, we projected the generated activation times onto the surface of the cardiac mesh. To compute the onset distance between the real and the predicted activation map (mean over 10 predicted maps from the β -CVAE), we considered arbitrarily in total seven onsets (5 on LV and 2 on RV). More precisely, for each pair of activation maps (real and generated), we computed the position of the first 5 activation times on the LV surface and the 2 earliest activation times on the RV. To ensure that the points within the map are distinct from each other, we impose a threshold distance of 15 mm. In Figure 5.9, we can visualise the onset pairs (reference/prediction) obtained for #Patient10. The larger spheres represent the reference points while the smaller ones are the prediction points. The same colour indicates a pair of onsets. For each of the two patients selected to evaluate of our model, we used the real raw signals (i.e., without any processing but nevertheless projected on our new torso model), to generate 10 predictions of cardiac activation.

Indice & location	Physical distance (mm)	Temporal distance (ms)	Predicted time (ms)	Reference time (ms)
#1 LV	13.76	47.4	48.5	1.0
#2 LV	5.77	35.6	37.7	2.1
#3 LV	11.4	23.3	29.8	6.5
#4 RV	21.2	33.7	40.3	6.6
#5 LV	14.3	32.4	39.5	7.1
#6 LV	6.06	22.9	30.2	7.3
#7 RV	30.1	30.8	40.3	9.4

Tab. 5.1: Temporal and physical distances between analogous onsets from predicted and real activation maps. The results presented are for #Patient12.

Figure 5.8 shows the results obtained by testing the model on #Patient12 data. It provides a visual comparison between the actual activation maps (a) and the mean of the 10 predictions (b). The std map (c) is used to report the variability of the estimation. Indeed, the areas corresponding to low std values indicate that the model has generated a stable activation map. On the contrary, the larger values designate where the model struggles to generate similar activation times. Looking at the map estimated by our model, we can see that the wave propagates through the same areas and in the same order as on the real map. Overall, the patterns of the estimated map correspond rather

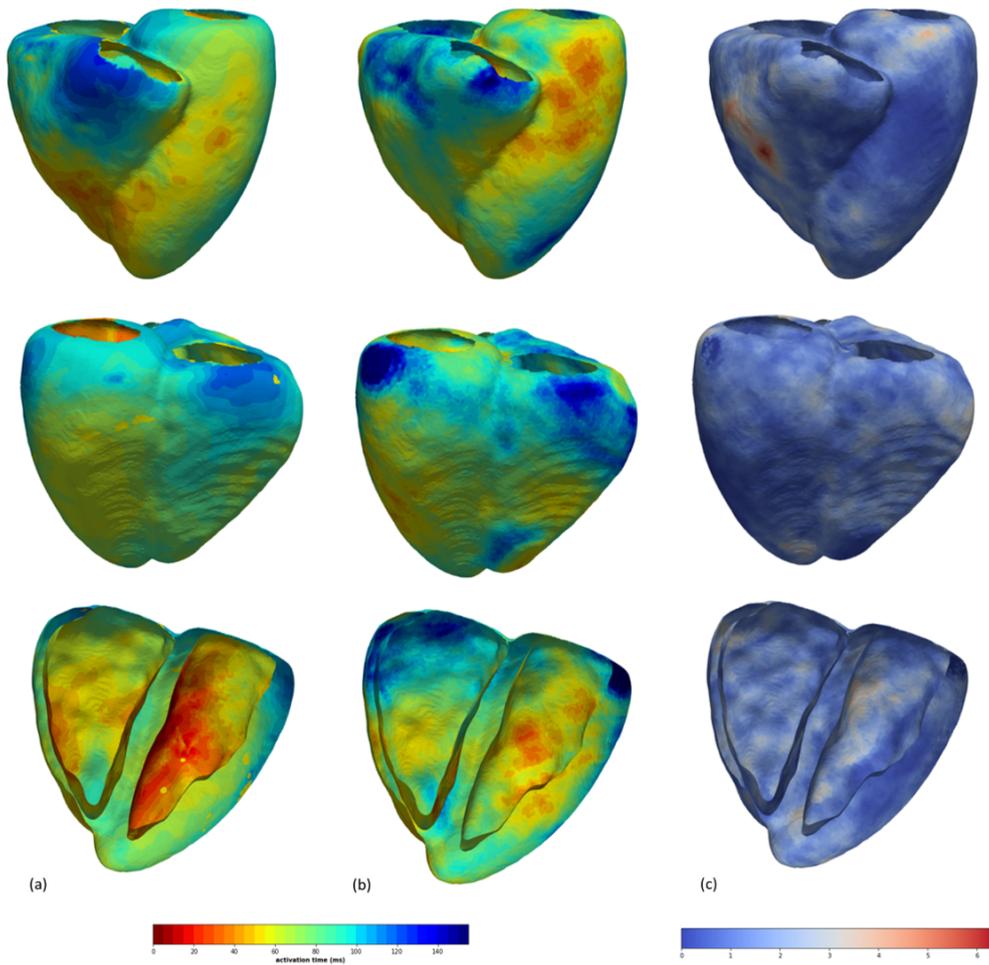


Fig. 5.8: Evaluation of estimated activation map from clinical BSP data by our proposed DL-based model. #Patient12: (a) ground truth (from CARTO) activation map; (b) mean activation map over 10 predictions, generated by our model; and (c) standard deviation map calculated over 10 prediction.

well to those of the real map. Although the patterns are well identified, their associated activation times do not correspond from one map to another. Furthermore, it can be noticed that the area of an activation zone is often larger in the real map than in the estimated map. The mean absolute error by testing the model on #Patient12 data is 27.6 ms. The results obtained in terms of physical and temporal distances between the 7 pairs of onsets are presented in Table 5.1. Each row corresponds to a pair of points. The mean euclidean distance for all pairs is 14.6 mm and the mean temporal distance 32.3 ms. The times of all the reference onsets are less than 10 ms while the earliest of the predicted times is about 30 ms. It is noticeable that all the times of the predictions are out of sync with the reals, this is not an exception.

Figure 5.10 illustrates the results obtained by testing the model with data from #Patient10. The model was effective in predicting patterns of early activation but has more difficulty in estimating areas of slow propagation. The mean absolute error is 20.5 ms.

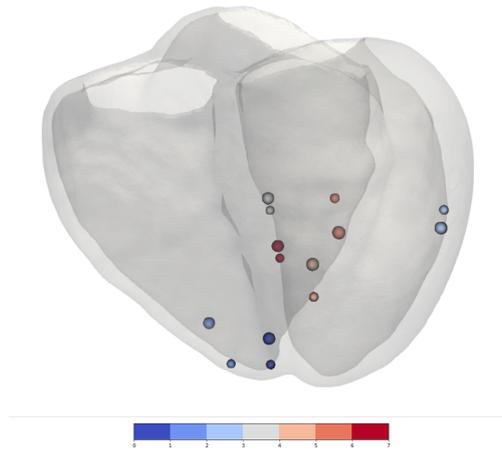


Fig. 5.9: Representation of the anatomy of #Patient10 and the 7 pairs of onsets. The larger spheres represent the reference onsets, the smaller ones correspond to the prediction onsets.

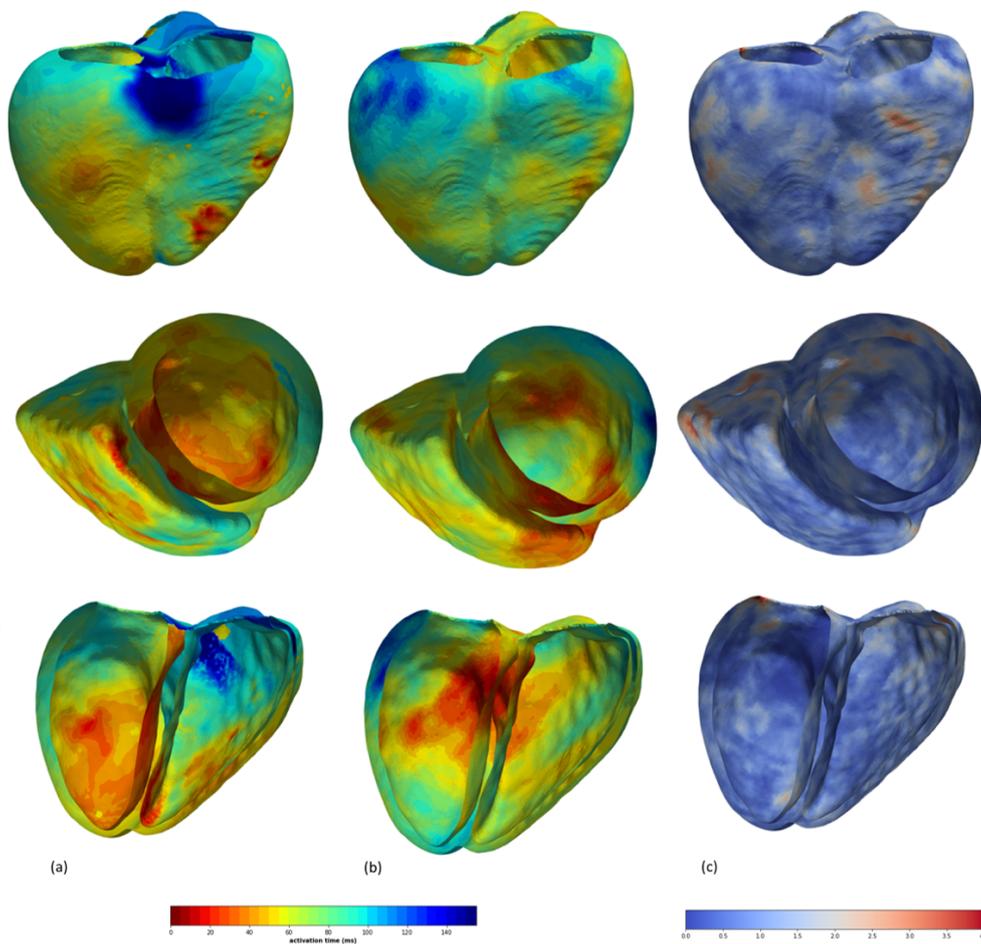


Fig. 5.10: Results obtained by testing the model on the anatomy of #Patient10

When reading Table 5.2, it can be seen that here, in contrast to the reference onsets of #Patient12 (see Table 5.1), the activation times of these points vary over a wider range. The minimum time is at 0 ms while the latest is at 27 ms. Moreover, here we observe that

Indice & location	Physical distance (mm)	Temporal distance (ms)	Predicted time (ms)	Reference time (ms)
#1 RV	9.02	56.8	56.8	0.00
#2 RV	15.3	47.3	54.5	7.21
#3 LV	5.80	4.36	20.6	16.2
#4 LV	7.88	25.2	43.0	17.8
#5 LV	10.3	35.3	54.6	19.3
#6 LV	10.3	3.27	20.3	23.6
#7 LV	5.41	27.0	54.1	27.1

Tab. 5.2: Temporal and physical distances between analogous onsets from predicted and real activation maps from #Patient10.

there are two exceptions (pairs of points #3 and #6) for which their activation times are of the same range. The activation time of reference point #1 is completely out of phase with the predicted one. Contrary to the results observed for #Patient12, here the time lag between baseline/predicted seems much more irregular. This may explain why our model has generated some areas of the map with great accuracy while some other areas have been poorly estimated. The mean physical distance is 9.15 mm and the mean temporal distance 28.4 ms.

5.4 Discussion

The proposed method has shown good effectiveness in predicting the correct direction of propagation as well as in identifying patterns of electrical activation of the heart. With an average distance of 11.9 mm for all cases, our model predicts the location of onsets reasonably satisfactory. The temporal distance between the onsets is however not optimal. The onsets showed a temporal discrepancy between the predicted and the real activation times, especially in some places. One explanation for this discrepancy could be the QRS duration in the real data, which theoretically seems too long. However, we believe that it is not essential to predict the exact activation time at a given point. The correct prediction of the pattern, the wave propagation and onset location are the most relevant features.

At this stage of our model's evaluation, we would like to expose the limitations we have to face and the possible perspectives to overcome these problems.

Firstly, we have noticed a significant difference between the signals we simulate and the real ones. Figure 5.11 shows real potentials (blue) and our adapted simulated signals (orange). (a) shows the signals at three different electrodes on the surface of the #Patient12 torso. On the left plot, the signals have a high correlation coefficient (cc). On the centre, are represented signals with median cc, while in the last plot are shown signals belonging to the 25th percentile of cc. (b) follows the same rules as (a)

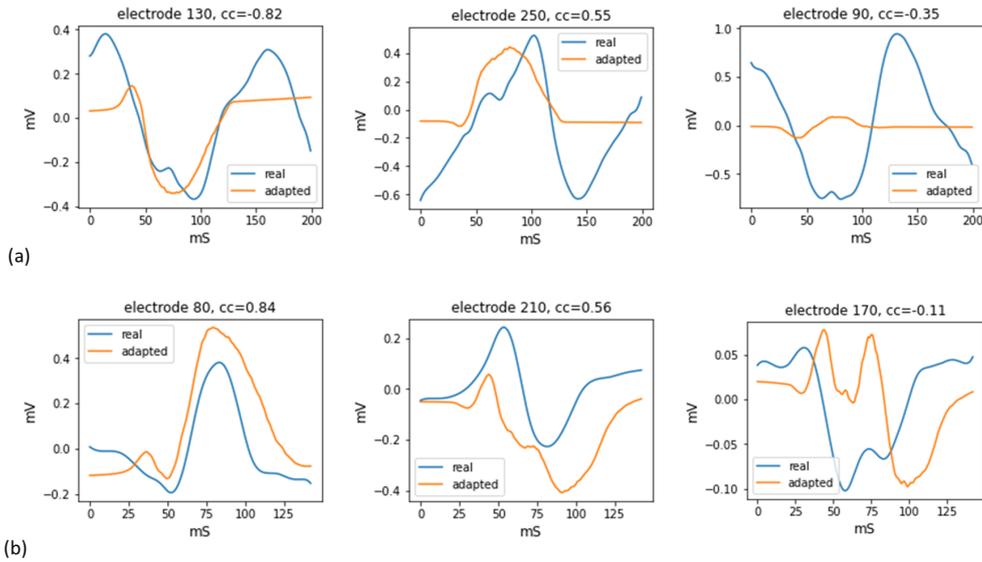


Fig. 5.11: Representation of the potentials at different electrodes of #Patient12 (a), #Patient10 (b). The blue curves represent the real signal, and the orange curve our adapted signal. The potentials represented in the first column belong to high correlated real-adapted signals (75th percentile of correlation coefficient). The second column, signals with median correlation, and in the last column are the 25th percentile of correlation signals.

but illustrates potentials at electrodes of #Patient10. On average on the two patients, only in 14.6% of the cases, the real and simulated signals pass the same number of times per zeros. The maximum amplitude of the real signals is on average 25% higher than the simulated ones. Furthermore, by observing the potentials on T_m (in the same way as illustrated in Figure 5.3), we have found that on the back of the torso there is no important propagation wave. In our simulations this is not the case. This problem is easily solved by applying a constraint to attenuate the potentials at these electrodes. To overcome these problems, an obvious solution would be to place even more emphasis on signal processing. Another solution that makes sense would be to extend the model to learn how to handle missing values (by removing some in the training phase) and during the test remove the electrodes where the similarities between real and simulated signals are very poor.

Secondly, the intra real data QRS, i.e., the QRS duration of the Carto data and the BSP for the same patient are very often significantly different. Sometimes the spread of activation in Carto data seems to be disordered, more precisely, we have cases where activation starts on the LV or RV followed by the epicardium. However, the activation does not end on the epicardium but on the LV or RV. In order to evaluate the model on clinical data of more patients, it will be necessary to establish a method to deal with these inconsistencies in the data.

A comparison between our predictions and the activation maps estimated by the CardioInsight technology can easily be made soon.

It would also be interesting to build a new training data set, using fewer patient anatomies for which we have BSPs and invasive activation maps usable for validation purposes. In order to have a consistent training data set, we plan to build a new synthetic database from the 6 anatomies of section 3. Ideally, some real cases should also be included in the training of the model. In this way, the method we have proposed can be tested on more patients data.

5.5 Conclusion

Our probabilistic generative model is able to generate cardiac activation maps for a new patient's geometry from non-invasive real BSP data and cardiac imaging. In the training phase, the model learns probabilistically the spatio-temporal and geometrical information from synthetic data. The generative part of our model, i.e., the decoder takes \bar{z} as well as the real non-invasive BSP and the cardiac propagation conduction velocity map as input data, and estimates the corresponding electrical cardiac activity. Evaluated on clinical data from two patients, the new formulation seems promising for reconstructing cardiac activation patterns and locating the position of onsets.

Our method provides a cardiac activation map in record time whereas most methods that have been found to be effective either require a long computation time or a reconstruction of the potential at the electrical surface. To the best of our knowledge, the evaluation we have proposed in this chapter is a first for a deep learning based model which evaluation has been performed qualitatively and quantitatively with real human invasive data.

Our novel framework meets several criteria that leads us to believe that it can be quickly integrated into clinical routine for non-invasive, accurate diagnostic and ECGI-based testing in large cohorts of arrhythmia patients.

Conclusions

Contents

6.1	Contributions of the Thesis	103
6.2	Perspectives	105
6.2.1	From 12-lead ECG	106
6.2.2	Interpretability of the latent space	107
6.2.3	A collaborative framework for ECGI clinical data processing	107

In the following sections, we summarise the methodological contributions of this thesis. Considering the current limitations, we propose research perspectives in the field of Deep learning and Electrocardiography involving our work.

6.1 Contributions of the Thesis

The use of Deep Learning to solve the inverse problem of Electrocardiography is the central focus of this thesis. By investigating how a model could rapidly generate a patient’s cardiac activation map from non-invasive data, i.e., the signals measured on the surface of its torso and its anatomical images, we conducted the following research:

In **Chapter 2**, we posed the problem in two dimensions. We introduced the variational autoencoder. Our multi-modal data are images and signals, i.e., data that contain temporal and spatial information, so we chose to use convolutional neural networks in our model. Indeed, the latter allows spatial information to be propagated through all layers of the network. We trained and tested two VAE independently. One was aimed at generating activation maps, the other at generating BSPs. These tests allowed us to understand how to choose the right hyperparameters for the VAE. We extended the VAE mono modality to the inverse problem of the ECGI. For this purpose, we proposed a β -CVAE, a VAE conditioned by non-invasive data (cardiac geometry and BSP), which once trained, was able to reconstruct the cardiac activation slice from the corresponding mask and signals with good accuracy. This work was the subject of a publication [Bacoyannis, 2019].

The CVAE presented in **Chapter 3** is an extension of the 2D formulation into 3D. The synthetic database was simulated considering only single onsets. The Eikonal model was

used to simulate the activation maps. The corresponding BSPs were calculated using the dipole formulation of the volume conductor theory and with the assumption that the thoracic domain is homogeneous and infinite. The proposed β -CVAE is capable of generating for a new unseen geometry, its activation map using the myocardial mask and the signals simulated on a 10×10 electrode Cartesian grid. This work was published in [Bacoyannis, 2021]

In **Chapter 4**, a multi-modal data processing framework including: invasive EGMs, CT images and torso signals was implemented. We also proposed a simplified geometric thoracic model. As clinical data is both rare and precious, we developed several processing methods to address each problem we faced in order to preserve as much of the data as possible and making it usable for research.

Finally, in the **Chapter 5**, the proposed 3D model was modified to be more adapted to real clinical data. The simulated data to produce the training data set was increased and considered multiple onsets. This allowed the database to be enriched. The model was trained on simulated data and evaluated on real clinical data. Our model allows for fast generation of activation maps from clinically-acquired BSPs. By probabilistically learning the spatio-temporal correlation and integrating anatomical information, the method goes beyond the current state-of-the-art, with comparison to invasive contact mapping data demonstrating accurate identification of activation patterns and a precise localisation of the onset of focal activation. This work will be presented at the EHRA 2022 conference.

This work leads to the following contributions:

- Spatio-temporal correlations: the convolutional CVAE model learns interactions in space and time between signals, while most ECGI methods solve each time step independently or use a temporal prior without spatial correlations.
- Imaging substrate: the correlation between the substrate from imaging (i.e. myocardial scar) and the signals is also learned, therefore we can seamlessly integrate any 3D image information in ECGI, while this is still difficult in the classical formulation.
- Data-driven regularisation: using a generative model should ensure smooth variations between similar cases and sampling from a low dimensional space should reduce the impact of noise; therefore this could alleviate the ill-posedness issue.
- Simplified personalised torso model: allowing us not only to accelerate the speed of simulation of BSPs, but also to reduce the computational time of the training of the DL-based model.

- Cardiac propagation conduction velocity map: using it as conditioning data in our model, after having specified constrain on it and adding it to the simulation pipeline, allows to better/more realistically reproduce the cardiac activation pattern.
- Body Surface Potential processing: applying it to our simulated BSP so that they are more consistent with real BSP, and integrate this processing into the training phase, allows our model to learn correlations that link activation maps to signals closer to reality.
- Fast computations: once trained, our DL method is very fast to evaluate.
- Volumetric predictions: this 3D convolutional approach generates volumetric activation maps within the whole myocardium, not limited to the epicardium and the endocardium.
- Processing framework for ECGI clinical data fusion: It consists on developped algorithms dedicated to the processing of: CT-scan images, EAM mapping and meshes and Body Surface Potentials. Other algorithms and techniques have been developed to integrate information from all modalities into a common reference space.
- Rich synthetic database: composed of 9900 activation maps and their corresponding BSPs. Three scenarios for simulating cardiac activation maps have been considered: one onset on the right ventricle, one onset on the left ventricle, three total onsets (1 on LV plus 2 on RV). All of the initialisation points were chosen randomly in accordance with the different scenarios.
- ECGI clinical database: composed of 11 patients. It consists of geometric data of the heart and torso, Body Surface Potentials and bipolar local activation times on the segmented cardiac mesh.
- Ready for use model: trained on 4 500 data, as soon as a new clinical data become available, it could generate its corresponding activation maps in 1 second by providing it only the cardiac geometry and the non-invasive BSP.

6.2 Perspectives

The method we proposed is directly related to the measured signals and therefore the maps generated by our β -CVAE are very sensitive to the BSPs it has been trained with. The more similar the BSPs used for training are to real signals, the better the performance

of our estimates in tests on clinical data. The real BSPs provided to us are from a commercialised technology, which means that filtering and processing that will never be disclosed to us have been done upstream. Based on what is known, i.e. the use of a reference electrode, we used the WCT technique by analogy, as the chosen reference electrode is not annotated in our data. Moreover, by observing the differences between our simulated signals after being corrected with WCT and the real signals, we observed that the latter never start at zero unlike ours. From this observation we chose to add a randomly generated low frequency ($0 - 0.5Hz$) sinusoidal noise for each electrode. Thus our signals are more in line with reality, but despite these processing they are still very different. In order to evaluate these differences, we calculated the correlation, the number of times the signal passes through zero, the amplitude under the curve. The values obtained for these metrics confirmed our suspicions. Data from patients who have undergone a catheter intervention are very rare and valuable.

6.2.1 From 12-lead ECG

In order to be able to keep them in the database we have built up, one idea would be to learn to reconstruct standard 12-lead ECGs from BSPs. As the clinical modalities for obtaining 12-lead ECGs and BSPs are not invasive, it is conceivable that a database could be obtained with these two measurements, combined with the torso geometry of the corresponding patient/volunteer. Thus, one could train a β -CVAE by taking as conditional data the geometry and the BSP, x being the 12-lead ECG. The aim is to learn to reconstruct the standard ECG. Once the model has been trained, we can estimate the 12-lead ECG corresponding to the BSP of our current data and thus constitute our new test data set. From the activation maps that we will simulate, we will simulate the corresponding 12-lead ECG to build a new training database.

This method would therefore allow us to estimate the activations of new patients from classic ECG and thus ideally to replace CardioInsight. Nowadays the number of electrodes and their positions to get the most informative signals can still be improved. One could also vary the number of electrodes and their positions on the basis of our electrode model. From these different combinations, we could simulate their corresponding BSP and use these signals as condition data in the test phase. Thus the closer the generated map is to reality, the more informative the BSP information will be.

The use of more complex models of cardiac electrophysiology and more detailed torso models would enable to evaluate their impact on the realism of the simulated signals.

6.2.2 Interpretability of the latent space

In this thesis, we did not explore the latent space. In contrast to another well-known dimension reduction technique, PCA, our CVAE model does not offer a simple and immediately accessible interpretability of its reduced dimensional space: the latent space. However, in our model, we have forced the latent representation to follow a normal distribution, which has the effect of pushing the various similarities that could be defined as 'clusters' to be as closely related as possible, grouping around a mean of zero. The β parameter we introduced in our CVAE intends to discover disentangled latent factors. To interpret the latent space, we would have to experiment with this parameter much more to find the optimal value for the observations of interest. By studying our latent space more deeply, we could find the most representative features in our data for each of the different pathologies on which our model will be trained. We did not explore this, as we only analysed ischemic cases. Collecting this information could be very interesting from a medical point of view and would also allow our model to evolve and improve.

6.2.3 A collaborative framework for ECGI clinical data processing

The clinical database built during this thesis has been notably achieved thanks to a mutualisation of the data processing between our team and the Carmen team of Inria Bordeaux. In the same continuity, it would be interesting and profitable for all to standardise the processing methods of these data, which have been proposed and developed separately. One objective would also be to automate as many of these methods as possible and integrate them into a workflow. Among the processing tasks we have identified, a solution should be proposed for the Carto data if possible. Specifically, there are many cases, where the QRS are of the order of 240 ms. For some cases, activation starts on the LV or RV followed by the epicardium. However, the activation does not end on the epicardium but on the LV or RV.

Collaboration with clinicians and with teams working on these Carto data would help to better understand the origin of this problem/observation and find a solution to solve it. Furthermore, by solving this, our generative model can be evaluated by testing it on the rest of the patients in our database who were never seen during training.

Bibliography

- [Alberts, 2002] B. Alberts, D. Bray, J. Lewis, et al. *Molecular Biology of the Cell*. 4th. Garland, 2002 (cit. on p. 11).
- [Aliev, 1996] R. Aliev and A. Panfilov. “A simple two-variable model of cardiac excitation”. In: *Chaos Solitons & Fractals* 7 (1996), pp. 293–301 (cit. on p. 14).
- [Aliot, 2009] Etienne M. Aliot, William G. Stevenson, Jesus Ma Almendral-Garrote, et al. “EHRA/HRS expert consensus on catheter ablation of ventricular arrhythmias”. English. In: *Europace* 11.6 (June 2009), pp. 771–817 (cit. on pp. 11, 12).
- [Arridge, 2019] Simon Arridge, Peter Maass, Ozan Öktem, and Carola-Bibiane Schönlieb. “Solving inverse problems using data-driven models”. In: *Acta Numerica* 28 (2019), pp. 1–174 (cit. on pp. 5, 38).
- [Arrieula, 2021] Andony Arrieula, Hubert Cochet, Pierre Jais, et al. “In-Silico Data Based Machine Learning Technique Predicts Premature Ventricular Contraction Origin Coordinates”. In: *Computing in Cardiology 2021-48th Computing in Cardiology Conference*. 2021 (cit. on p. 84).
- [Attene, 2010] Marco Attene. “A lightweight approach to repairing digitized polygon meshes”. In: *The Visual Computer* 26 (Nov. 2010), pp. 1393–1406 (cit. on p. 67).
- [Bacoyannis, 2019] Tania Bacoyannis, Julian Krebs, Nicolas Cedilnik, Hubert Cochet, and Maxime Sermesant. “Deep Learning Formulation of ECGI for Data-Driven Integration of Spatiotemporal Correlations and Imaging Information”. In: *Functional Imaging and Modeling of the Heart*. Ed. by Yves Coudière, Valéry Ozenne, Edward Vigmond, and Nejib Zemzemi. Cham: Springer International Publishing, 2019, pp. 20–28 (cit. on pp. 38, 45, 103).
- [Bacoyannis, 2021] Tania Bacoyannis, Buntheng Ly, Nicolas Cedilnik, Hubert Cochet, and Maxime Sermesant. “Deep learning formulation of electrocardiographic imaging integrating image and signal information with data-driven regularization”. In: *EP Europace* 23.Supplement₁ (Mar. 2021), pp. i55–i62 (cit. on pp. 57, 81, 84, 87, 92, 104).

- [Barr, 1977] R. C. Barr, M. Ramsey, and M. S. Spach. “Relating Epicardial to Body Surface Potential Distributions by Means of Transfer Coefficients Based on Geometry Measurements”. In: *IEEE Transactions on Biomedical Engineering* BME-24.1 (1977), pp. 1–11 (cit. on pp. 15, 37).
- [Bear, 2018a] Laura R Bear, Y Serinagaoglu Dogrusoz, Jana Svehliková, et al. “Effects of ecg signal processing on the inverse problem of electrocardiography”. In: *2018 Computing in Cardiology Conference (CinC)*. Vol. 45. IEEE. 2018, pp. 1–4 (cit. on pp. 17, 38).
- [Bear, 2018b] Laura R Bear, Ian J LeGrice, Gregory B Sands, et al. “How accurate is inverse electrocardiographic mapping? A systematic in vivo evaluation”. In: *Circulation: Arrhythmia and Electrophysiology* 11.5 (2018), e006108 (cit. on p. 16).
- [Cedilnik, 2018] Nicolas Cedilnik, Josselin Duchateau, Rémi Dubois, et al. “Fast Personalized Electrophysiological Models from CT Images for Ventricular Tachycardia Ablation Planning”. In: *EP-Europace* 20 (Nov. 2018) (cit. on pp. 15, 30, 43, 44, 47, 68, 89).
- [Cedilnik, 2019] Nicolas Cedilnik, Josselin Duchateau, Frédéric Sacher, et al. “Fully Automated Electrophysiological Model Personalisation Framework from CT Imaging”. In: *Functional Imaging and Modeling of the Heart*. Ed. by Yves Coudière, Valéry Ozenne, Edward Vigmond, and Nejib Zenzemi. Cham: Springer International Publishing, 2019, pp. 325–333 (cit. on pp. 38, 45, 54, 87).
- [Chamorro-Servent, 2017] Judit Chamorro-Servent, Rémi Dubois, Mark Potse, and Yves Coudière. “Improving the Spatial Solution of Electrocardiographic Imaging: A New Regularization Parameter Choice Technique for the Tikhonov Method”. In: *Functional Imaging and Modelling of the Heart*. Ed. by Mihaela Pop and Graham A Wright. Cham: Springer International Publishing, 2017, pp. 289–300 (cit. on pp. 16, 27, 37).
- [Chávez, 2015] Carlos Eduardo Chávez, Nejib Zenzemi, Yves Coudière, Felipe Alonso-Atienza, and Diego Álvarez. “Inverse Problem of Electrocardiography: Estimating the Location of Cardiac Ischemia in a 3D Realistic Geometry”. In: *Functional Imaging and Modeling of the Heart*. Ed. by Hans van Assen, Peter Bovendeerd, and Tammo Delhaas. Cham: Springer International Publishing, 2015, pp. 393–401 (cit. on p. 43).
- [Chavez, 2017] C.E. Chavez, Nejib Zenzemi, Yves Coudière, Felipe Alonso-Atienza, and Diego Alvarez. “Inverse Problem of Electrocardiography: estimating the location of cardiac ischemia in a 3D geometry”. In: *Functional Imaging and modelling of the heart (FIMH2015)*. Vol. 9126. Springer, 2017 (cit. on p. 27).
- [Cluitmans, 2018] Matthijs Cluitmans, Dana H. Brooks, Rob MacLeod, et al. “Validation and Opportunities of Electrocardiographic Imaging: From Technical Achievements to Clinical Applications”. In: *Frontiers in Physiology* 9 (2018), p. 1305 (cit. on pp. 17, 38, 83).

- [Danielsson, 1980] Per-Erik Danielsson. “Euclidean distance mapping”. In: *Computer Graphics and Image Processing* 14.3 (1980), pp. 227–248 (cit. on p. 64).
- [Dapogny, 2014] Charles Dapogny, Cécile Dobrzynski, and Pascal Frey. “Three-dimensional adaptive domain remeshing, implicit domain meshing, and applications to free and moving boundary problems”. In: *Journal of Computational Physics* (Apr. 2014) (cit. on pp. 55, 61).
- [Desrues, 2021] Gaëtan Desrues, Delphine Feuerstein, Thierry Legay, Serge Cazeau, and Maxime Sermesant. “Personal-by-design: a 3D Electromechanical Model of the Heart Tailored for Personalisation”. In: *Functional Imaging and Modeling of the Heart*. Stanford, CA, United States, June 2021 (cit. on p. 90).
- [Fitzhugh, 1961] R. Fitzhugh. “Impulses and physiological states in theoretical models of nerve membrane”. In: *Biophysical Journal* 1 (1961), pp. 455–466 (cit. on p. 14).
- [Franzone, 1990] Piero Colli Franzone, Lucia Guerri, and Sergio Rovida. “Wavefront propagation in an activation model of the anisotropic cardiac tissue: asymptotic analysis and numerical simulations”. In: *Journal of Mathematical Biology* 28 (1990), pp. 121–176 (cit. on p. 15).
- [Gardner, 1985] Phyllis I. Gardner, Philip C. Ursell, John J. Fenoglio, and Andrew L. Wit. “Electrophysiologic and anatomic basis for fractionated electrograms recorded from healed myocardial infarcts.” In: *Circulation* 72 3 (1985), pp. 596–611 (cit. on p. 11).
- [Ghimire, 2018] Sandesh Ghimire, Jwala Dhamala, Prashna Kumar Gyawali, et al. “Generative Modeling and Inverse Imaging of Cardiac Transmembrane Potential”. In: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*. Cham: Springer, 2018, pp. 508–516 (cit. on pp. 17, 23, 26, 38, 84).
- [Ghimire, 2019] Sandesh Ghimire, Prashna Kumar Gyawali, Jwala Dhamala, et al. “Improving Generalization of Deep Networks for Inverse Reconstruction of Image Sequences”. In: *Information Processing in Medical Imaging*. Ed. by Albert C. S. Chung, James C. Gee, Paul A. Yushkevich, and Siqi Bao. Cham: Springer International Publishing, 2019, pp. 153–166 (cit. on p. 17).
- [Ghosh, 2009] Subham Ghosh and Yoram Rudy. “Application of l1-norm regularization to epicardial potential solution of the inverse electrocardiography problem”. In: *Annals of biomedical engineering* 37.5 (2009), pp. 902–912 (cit. on p. 16).
- [Giffard-Roisin, 2017] Sophie Giffard-Roisin, Thomas Jackson, Lauren Fovargue, et al. “Non-Invasive Personalisation of a Cardiac Electrophysiology Model from Body Surface Potential Mapping”. In: *IEEE Transactions on Biomedical Engineering*. IEEE Transactions on Biomedical Engineering 64.9 (Sept. 2017), pp. 2206–2218 (cit. on pp. 30, 44, 90).

- [Giffard-Roisin, 2018] Sophie Giffard-Roisin, Hervé Delingette, Thomas Jackson, et al. “Transfer Learning from Simulations on a Reference Anatomy for ECGI in Personalised Cardiac Resynchronization Therapy”. In: *IEEE Transactions on Biomedical Engineering* 20 (2018) (cit. on pp. 17, 23, 34, 48, 83).
- [Greensite, 1998] Fred Greensite and Geertjan Huiskamp. “An improved method for estimating epicardial potentials from the body surface”. In: *IEEE Transactions on Biomedical Engineering* 45.1 (1998), pp. 98–104 (cit. on p. 16).
- [Hammernik, 2018] Kerstin Hammernik, Teresa Klatzer, Erich Kobler, et al. “Learning a variational network for reconstruction of accelerated MRI data”. In: *Magnetic Resonance in Medicine* 79.6 (2018), pp. 3055–3071 (cit. on pp. 5, 26, 38).
- [Han, 2015] Chengzong Han, Steven M Pogwizd, Long Yu, et al. “Imaging cardiac activation sequence during ventricular tachycardia in a canine model of nonischemic heart failure”. In: *American Journal of Physiology-Heart and Circulatory Physiology* 308.2 (2015), H108–H114 (cit. on p. 16).
- [He, 2003] Bin He, Guanglin Li, and Xin Zhang. “Noninvasive imaging of cardiac transmembrane potentials within three-dimensional myocardium by means of a realistic geometry anisotropic heart model”. In: *IEEE Transactions on Biomedical Engineering* 50.10 (2003), pp. 1190–1202 (cit. on p. 16).
- [Higgins, 2017] Irina Higgins, Loic Matthey, Arka Pal, et al. “ β -VAE: LEARNING BASIC VISUAL CONCEPTS WITH A CONSTRAINED VARIATIONAL FRAMEWORK”. In: *International Conference on Learning Representations*. 2017 (cit. on pp. 22, 29, 40, 92).
- [HODGKIN, 1952] AL HODGKIN and AF HUXLEY. “A quantitative description of membrane current and its application to conduction and excitation in nerve”. In: *The Journal of physiology* 117.4 (1952), pp. 500–544 (cit. on p. 8).
- [Huiskamp, 1989] Geertjan Huiskamp and Adriaan van Oosterom. “Tailored versus realistic geometry in the inverse problem of electrocardiography”. In: *IEEE transactions on biomedical engineering* 36.8 (1989), pp. 827–835 (cit. on p. 16).
- [Jacquemet, 2011] Vincent Jacquemet. “An eikonal-diffusion solver and its application to the interpolation and the simulation of reentrant cardiac activations”. In: *Computer methods and programs in biomedicine* 108 (June 2011), pp. 548–58 (cit. on p. 15).
- [Jia, 2018] Shuman Jia, Antoine Despinasse, Zihao Wang, et al. “Automatically Segmenting the Left Atrium from Cardiac Images Using Successive 3D U-Nets and a Contour Loss”. In: *STACOM: Statistical Atlases and Computational Models of the Heart. Atrial Segmentation and LV Quantification Challenges*. Vol. 11395. LNCS. Granada, Spain, Sept. 2018, pp. 221–229 (cit. on p. 54).

- [Jiang, 2009] Yuan Jiang, C. Qian, R. Hanna, D. Farina, and O. Dössel. “Optimization of the Electrode Positions of Multichannel ECG for the Reconstruction of Ischemic Areas by Solving the Inverse Electrocardiographic Problem”. In: 2009 (cit. on p. 13).
- [Jiang, 2010] Ming Feng Jiang, Chen Qun Wang, Wen Qing Huang, and Xia Cai. “The application of support vector regression method for solving the inverse ecg problem”. In: *Advanced Materials Research*. Vol. 108. Trans Tech Publ. 2010, pp. 828–833 (cit. on pp. 22, 83).
- [Jiang, 2011] Mingfeng Jiang, Lingyan Zhu, Yaming Wang, et al. “Application of kernel principal component analysis and support vector regression for reconstruction of cardiac transmembrane potentials”. In: *Physics in Medicine and Biology* 56.6 (2011), pp. 1727–1742 (cit. on p. 22).
- [Jiang, 2013] Mingfeng Jiang, Yaming Wang, Ling Xia, et al. “The combination of self-organizing feature maps and support vector regression for solving the inverse ECG problem”. In: *Computers Mathematics with Applications* 66.10 (2013). ICNC-FSKD 2012, pp. 1981–1990 (cit. on p. 22).
- [Karoui, 2018] Amel Karoui, Laura Bear, Pauline Migerditichan, and Nejib Zemzemi. “Evaluation of Fifteen Algorithms for the Resolution of the Electrocardiography Imaging Inverse Problem Using ex-vivo and in-silico Data”. In: *Frontiers in Physiology* 9 (2018), p. 1708 (cit. on pp. 17, 38, 83).
- [Karoui, 2019] Amel Karoui, Mostafa Bendahmane, and Nejib Zemzemi. “A Spatial Adaptation of the Time Delay Neural Network for Solving ECGI Inverse Problem”. In: *International Conference on Functional Imaging and Modeling of the Heart*. Springer. 2019, pp. 94–102 (cit. on pp. 17, 23).
- [Keener, 1991] James P. Keener. “The effects of discrete gap junction coupling on propagation in myocardium”. In: *Journal of Theoretical Biology* 148.1 (1991), pp. 49–82 (cit. on p. 15).
- [Kingma, 2014a] Diederik P. Kingma and Jimmy Ba. *Adam: A Method for Stochastic Optimization*. cite arxiv:1412.6980Comment: Published as a conference paper at the 3rd International Conference for Learning Representations, San Diego, 2015. 2014 (cit. on pp. 46, 93).
- [Kingma, 2014b] Diederik P. Kingma and Max Welling. “Auto-Encoding Variational Bayes”. In: *In: Proceedings of the International Conference on Learning Representations (ICLR) (2014)* (cit. on pp. 20, 21, 27–29, 39, 91, 92).
- [Kingma, 2014c] Durk P Kingma, Shakir Mohamed, Danilo Jimenez Rezende, and Max Welling. “Semi-supervised Learning with Deep Generative Models”. In: *Advances in Neural Information Processing Systems* 27. Ed. by Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger. Curran Associates, Inc., 2014, pp. 3581–3589 (cit. on pp. 22, 29, 39, 91).

- [Kohonen, 2003] Teuvo Kohonen. “Self-organized maps of sensory events”. In: *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 361.1807 (2003), pp. 1177–1186 (cit. on p. 23).
- [Lecun, 1998] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner. “Gradient-based learning applied to document recognition”. In: *Proceedings of the IEEE* 86.11 (1998), pp. 2278–2324 (cit. on p. 19).
- [LeCun, 2015] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. “Deep Learning”. In: *Nature* 521.7553 (2015), pp. 436–444 (cit. on p. 5).
- [Lines, 2003] GT Lines, ML Buist, P Grottum, et al. “Mathematical models and numerical methods for the forward problem in cardiac electrophysiology”. In: *Computing and Visualization in Science* 5.4 (2003), pp. 215–239 (cit. on p. 15).
- [Litjens, 2017] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, et al. “A survey on deep learning in medical image analysis”. In: *Medical Image Analysis* 42 (2017), pp. 60–88 (cit. on p. 5).
- [Liu, 2006] Zhongming Liu, Chenguang Liu, and Bin He. “Noninvasive reconstruction of three-dimensional ventricular activation sequence from the inverse solution of distributed equivalent current density”. In: *IEEE transactions on medical imaging* 25.10 (2006), pp. 1307–1318 (cit. on p. 16).
- [Lorensen, 1987] William E. Lorensen and Harvey E. Cline. “Marching cubes: A high resolution 3D surface construction algorithm.” In: *SIGGRAPH*. Ed. by Maureen C. Stone. ACM, 1987, pp. 163–169 (cit. on p. 63).
- [Malik, 2018] Avinash Malik, Tommy Peng, and Mark L Trew. “A machine learning approach to reconstruction of heart surface potentials from body surface potentials”. In: *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE. 2018, pp. 4828–4831 (cit. on p. 23).
- [Malmivuo, 1995] Jaakko Malmivuo and Robert Plonsey. *Bioelectromagnetism - Principles and Applications of Bioelectric and Biomagnetic Fields*. Oct. 1995 (cit. on p. 9).
- [Milanič, 2014] Matija Milanič, Vojko Jazbinšek, Robert S MacLeod, Dana H Brooks, et al. “Assessment of regularization techniques for electrocardiographic imaging”. In: *Journal of electrocardiology* 47.1 (2014), pp. 20–28 (cit. on p. 17).
- [Missel, 2020] Ryan Missel, Prashanna K. Gyawali, Jaideep Vitthal Murkute, et al. “A hybrid machine learning approach to localizing the origin of ventricular tachycardia using 12-lead electrocardiograms”. In: *Computers in Biology and Medicine* 126 (2020), p. 104013 (cit. on p. 84).
- [Mitchell, 2003] Colleen C Mitchell and David G Schaeffer. “A two-current model for the dynamics of cardiac membrane”. en. In: *Bulletin of Mathematical Biology* 65.5 (Sept. 2003), pp. 767–793 (cit. on pp. 14, 44).

- [Pereira, 2020] Helder Pereira, Steven Niederer, and Christopher A Rinaldi. “Electrocardiographic imaging for cardiac arrhythmias and resynchronization therapy”. In: *EP Europace* (Aug. 2020). euaa165. eprint: <https://academic.oup.com/europace/advance-article-pdf/doi/10.1093/europace/euaa165/33570218/euaa165.pdf> (cit. on pp. 15, 37).
- [Potse, 2006] M. Potse, B. Dube, J. Richer, A. Vinet, and R. M. Gulrajani. “A Comparison of Monodomain and Bidomain Reaction-Diffusion Models for Action Potential Propagation in the Human Heart”. In: *IEEE Transactions on Biomedical Engineering* 53.12 (Dec. 2006), pp. 2425–2435 (cit. on pp. 15, 43).
- [Pullan, 2004] A. Pullan, L. Cheng, M. Nash, C. Bradley, and D. Paterson. “Noninvasive Electrical Imaging of the Heart: Theory and Model Development”. In: *Annals of Biomedical Engineering* 29 (2004), pp. 817–836 (cit. on p. 16).
- [Ramanathan, 2001] Charu Ramanathan and Yoram Rudy. “Electrocardiographic Imaging: Effect of Torso Inhomogeneities on Noninvasive Reconstruction of Epicardial Potentials, Electrograms, and Isochrones”. In: *Journal of cardiovascular electrophysiology* 12 (Mar. 2001), pp. 241–52 (cit. on pp. 16, 27, 37).
- [Ringwala, 2019] Sukit Ringwala and Rishi Arora. “6 - Fundamentals of Intracardiac Mapping”. In: *Catheter Ablation of Cardiac Arrhythmias (Fourth Edition)*. Ed. by Shoen K. Stephen Huang and John M. Miller. Fourth Edition. Philadelphia: Elsevier, 2019, 76–100.e2 (cit. on p. 11).
- [Schroeder, 2006] Will Schroeder, Ken Martin, and Bill Lorensen. *The Visualization Toolkit—An Object-Oriented Approach To 3D Graphics*. Fourth. Kitware, Inc., 2006 (cit. on p. 54).
- [Schuler, 2017] Steffen Schuler, Danila Potyagaylo, and Olaf Dössel. “ECG imaging of simulated atrial fibrillation: Imposing epi-endocardial similarity facilitates the reconstruction of transmembrane voltages”. In: *2017 Computing in Cardiology (CinC)*. 2017, pp. 1–4 (cit. on p. 16).
- [Serinagaoglu, 2006] Yesim Serinagaoglu, Dana H Brooks, and Robert S MacLeod. “Improved performance of bayesian solutions for inverse electrocardiography using multiple information sources”. In: *IEEE Transactions on Biomedical Engineering* 53.10 (2006), pp. 2024–2034 (cit. on p. 16).
- [Sermesant, 2005] Maxime Sermesant, Yves Coudière, Valérie Moreau-Villéger, et al. “A fast-marching approach to cardiac electrophysiology simulation for XMR interventional imaging”. In: *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2005* (2005), pp. 607–615 (cit. on pp. 30, 43).
- [Shou, 2007] Guofa Shou, Ling Xia, Mingfeng Jiang, Feng Liu, and Stuart Crozier. “Forward and Inverse Solutions of Electrocardiography Problem Using an Adaptive BEM Method”. In: vol. 4466. June 2007, pp. 290–299 (cit. on p. 16).

- [Sohn, 2015] Kihyuk Sohn, Honglak Lee, and Xinchen Yan. “Learning Structured Output Representation using Deep Conditional Generative Models”. In: *Advances in Neural Information Processing Systems*. Ed. by C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett. Vol. 28. Curran Associates, Inc., 2015, pp. 3483–3491 (cit. on pp. 22, 39).
- [Soto-Iglesias, 2016] David Soto-Iglesias, Constantine Butakoff, David Andreu, et al. “Integration of electro-anatomical and imaging data of the left ventricle: An evaluation framework”. In: *Medical Image Analysis* 32 (2016), pp. 131–144 (cit. on pp. 12, 89).
- [Stenroos, 2008] Matti Stenroos and Jens Haueisen. “Boundary Element Computations in the Forward and Inverse Problems of Electrocardiography: Comparison of Collocation and Galerkin Weightings”. In: *Biomedical Engineering, IEEE Transactions on* 55 (Oct. 2008), pp. 2124–2133 (cit. on pp. 16, 27).
- [Stevenson, 2005] William Stevenson and Kyoko Soejima. “Recording Techniques for Clinical Electrophysiology”. In: *Journal of cardiovascular electrophysiology* 16 (Oct. 2005), pp. 1017–22 (cit. on pp. 11, 12).
- [Toloubidokhti, 2021] Maryam Toloubidokhti, Prashnna K Gyawali, Omar A Gharbia, et al. “Deep Adaptive Electrocardiographic Imaging with Generative Forward Model for Error Reduction”. In: *International Conference on Functional Imaging and Modeling of the Heart*. Springer. 2021, pp. 471–481 (cit. on p. 84).
- [Tusscher, 2004] K. H. W. J. ten Tusscher, D. Noble, P. J. Noble, and A. V. Panfilov. “A model for human ventricular tissue”. In: *American Journal of Physiology-Heart and Circulatory Physiology* 286.4 (2004). PMID: 14656705, H1573–H1589. eprint: <https://doi.org/10.1152/ajpheart.00794.2003> (cit. on p. 14).
- [Valette, 2004] Sébastien Valette and Jean-Marc Chassery. “Approximated Centroidal Voronoi Diagrams for Uniform Polygonal Mesh Coarsening”. In: *Computer Graphics Forum* 23.3 (2004). Eurographics 2004 proceedings, pp. 381–389 (cit. on p. 64).
- [Valette, 2008] Sebastien Valette, Jean Marc Chassery, and Remy Prost. “Generic Remeshing of 3D Triangular Meshes with Metric-Dependent Discrete Voronoi Diagrams”. In: *IEEE Transactions on Visualization and Computer Graphics* 14.2 (2008), pp. 369–381 (cit. on pp. 64, 87).
- [Van Oosterom, 2005] A Van Oosterom and V Jacquemet. “A parameterized description of transmembrane potentials used in forward and inverse procedures”. In: *Folia Cardiologica* 12.suppl D (2005), p. 111 (cit. on pp. 17, 38).
- [Vapnik, 1999] V.N. Vapnik. “An overview of statistical learning theory”. In: *IEEE Transactions on Neural Networks* 10.5 (1999), pp. 988–999 (cit. on pp. 22, 83).
- [Virtanen, 2020] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, et al. “SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python”. In: *Nature Methods* 17 (2020), pp. 261–272 (cit. on p. 68).

- [Wallman, 2012] M. Wallman, N. P. Smith, and B. Rodriguez. “A Comparative Study of Graph-Based, Eikonal, and Monodomain Simulations for the Estimation of Cardiac Activation Times”. In: *IEEE Transactions on Biomedical Engineering* 59.6 (2012), pp. 1739–1748 (cit. on p. 47).
- [Wang, 2006] Yong Wang and Yoram Rudy. “Application of the Method of Fundamental Solutions to Potential-based Inverse Electrocardiography”. In: *Annals of Biomedical Engineering* 34 (2006), pp. 1272–1288 (cit. on p. 16).
- [Wang, 2009a] Linwei Wang, H. Zhang, Ken C. L. Wong, H. Liu, and P. Shi. “Electrocardiographic simulation on personalised heart-torso structures using coupled meshfree-BEM platform”. In: *Int. J. Funct. Informatics Pers. Medicine* 2 (2009), pp. 175–200 (cit. on p. 16, 37).
- [Wang, 2009b] Linwei Wang, Heye Zhang, Ken CL Wong, Huafeng Liu, and Pengcheng Shi. “Physiological-model-constrained noninvasive reconstruction of volumetric myocardial transmembrane potentials”. In: *IEEE Transactions on Biomedical Engineering* 57.2 (2009), pp. 296–315 (cit. on p. 16).
- [Wang, 2010] Dafang Wang, Robert Kirby, and Chris Johnson. “Resolution Strategies for the Finite-Element-Based Solution of the ECG Inverse Problem”. In: *Biomedical Engineering, IEEE Transactions on* 57 (Mar. 2010), pp. 220–237 (cit. on p. 16).
- [Wang, 2019] L. Wang, Weiqi Wu, Yiping Chen, and Changhua Liu. “A New ISTANet Solution to Inverse Problem of Electrocardiology”. In: 2019 (cit. on p. 23).
- [Yedder, 2020] Hanene Ben Yedder, Ben Cardoen, and G. Hamarneh. “Deep learning for biomedical image reconstruction: a survey”. In: *Artificial Intelligence Review* 54 (2020), pp. 215–251 (cit. on p. 5).
- [Zemzemi, 2012] Nejib Zemzemi, Simon Labarthe, Rémi Dubois, and Yves Coudière. “From Body Surface Potential to Activation Maps on the Atria: a Machine Learning Technique”. In: *CINC - Computing in Cardiology 2012*. Vol. 39. Computing in Cardiology 2012, Krakow, Poland, Sept. 2012, pp. 125–128 (cit. on p. 23).
- [Zemzemi, 2015] Nejib Zemzemi, Cecile Dobrzynski, Laura Bear, et al. “Effect of the torso conductivity heterogeneities on the ECGI inverse problem solution”. In: *Computing in cardiology*. Nice, France, Sept. 2015 (cit. on pp. 16, 27, 37).
- [Zipes, 2018] Douglas P. Zipes, José Jalife, and William G. Stevenson. “Preface”. In: *Cardiac Electrophysiology: From Cell to Bedside (Seventh Edition)*. Ed. by Douglas P. Zipes, José Jalife, and William G. Stevenson. Seventh Edition. Elsevier, 2018, p. xix (cit. on p. 8).

