Phase-Independent Latent Representation for Cardiac Shape Analysis

Josquin Harrison[1], Marco Lorenzi[1], Benoit Legghe[2,3], Xavier Iriart[2,3], Hubert Cochet[2,3], and Maxime Sermesant[1,2(✉)]

[1] Inria, Université Côte d'Azur, Nice, France
maxime.sermesant@inria.fr
[2] IHU Liryc, Université de Bordeaux, Bordeaux, France
[3] Bordeaux University Hospital, Bordeaux, France

Abstract. Atrial fibrillation (AF) is a complex cardiac disease impacting an ever-growing population and increases 6-fold the risk of thrombus formation. However, image based bio-markers to predict thrombosis in presence of AF are not well known. This lack of knowledge comes from the difficulty to analyse and compare the shape of the Left Atrium (LA) as well as the insufficiency of data that limits the complexity of models we can use. Conducting data analysis in cardiology exacerbates the small dataset problem because the heart cycle renders impossible to compare images taken at systole and diastole time. To address these issues, we first propose a graph representation of the LA, to focus on the impact of pulmonary veins (PV) and LA Appendage (LAA) positions, giving a simple object easy to analyse. Secondly, we propose a meta-learning framework for heterogeneous datasets based on the consistent representation of each dataset in a common latent space. We show that such a model is analogous to a meta-classifier, where each dataset is characterised by specific projection in a common latent space, while sharing the same separating boundary. We apply this model to the graph representation of the LA and interpret the model to give novel time-dependant bio-markers related to PV and LAA configurations for the prediction of thrombosis.

Keywords: Shape analysis · Atrial fibrillation · Thrombosis · Graph representation · Latent space model · Multi-task learning · Meta-learning

1 Introduction

Atrial fibrillation (AF) is a complex cardiac disease impacting an ever-growing population, creating a hemodynamic environment prone to clot formation and ischemic stroke. The stratification of stroke risk in AF has significant clinical implications for the management of anticoagulation, which was shown to effectively limit the occurrence of strokes but at the cost of increased risk of bleeding. The net benefit of introducing preventive anticoagulation is currently estimated

by computing a score based on patient demographics, clinical condition and past history, but the ability of this score to accurately assess stroke risk in AF patients is largely suboptimal [8]. Thus, to better understand AF, important studies were conducted on LA hemodynamics [12] or LAA morphology [13]. Indeed, as the immense majority of clots occur in the LAA which is known to show high inter-individual variability, series of studies have focused on characterizing the LAA shape, demonstrating moderate association with stroke risk [3]. However, those studies are mostly qualitative, tools available to clinician for decision making are still limited and while imaging data availability is growing, it is under-utilised to quantitatively explore novel image-based bio-markers. An important hurdle making it difficult to use is its heterogeneity. In cardiology, images are taken at different times in the heart cycle, using different imaging systems, and for different tasks. Developing analysis methods allowing to integrate information across heterogeneous datasets to enable statistical studies is therefore a necessity.

In this work we follow this idea. Inspired by [1], we propose a multi-channel formulation to merge multiple heterogeneous datasets into a common latent representation. This framework realises a combination of Multi-task Learning (MTL) [9] and Meta Learning (ML) [4]: the common representation across datasets imposed in the latent space induces homogeneity across latent projections while enriching the amount of data provided to train a classifier for automated image-based diagnosis. In addition, the learned latent distribution of the joint representation is a meta parameter that constitutes an excellent prior for future datasets as it is robust to multiple datasets sharing a common task.

We propose to use such a model to explore underlying links between Thrombosis and Pulmonary Veins (PV), Appendage (LAA) positions and orientations in the Left Atrium. In the following sections we propose a lightweight graph representation of the LA to focus on PV and LAA positions and formulate the classification methodology within a supervised framework where the common representation is improved in terms of Kullback-Leibler divergence. We finally apply the model to the joint analysis of LA Graphs where the data is split in systole and diastole subsets. This constitutes a multi-label classification problem across datasets where the labels are consistent but the dataset is heterogeneous.

2   Methodology

2.1   Pre-processing Pipeline

Our study was performed on 3D Computational Tomography scans (CT-Scans) along with clinical data from 107 patients suffering atrial fibrillation, of which 64 are labelled Thrombus positive, a composite criterium composed of a detection of LAA thrombus on CT scan and/or past history embolism. In particular, our database is composed of 50 patients in systole (of which 27 are Thrombus positive) and 57 patients in diastole (of which 37 are Thrombus positive). Cardiac segmentation was first conducted automatically with a 3D U-Net neural network as proposed by [5], and then hand-corrected by experts. All the data was acquired by the Bordeaux University Hospital.

From the available segmentation masks we use the open-source package `MMG`[1] for meshing the shapes. First we apply a marching cube algorithm, giving a rough meshing of the surface, the mmg3d algorithm is then applied with specific parameters to keep the number of triangles under 2500.

2.2 Graph Representation of the LA

To study the impact of PVs and LAA geometry in clot formation, we propose to represent the LA as a graph similar to its centre-line. To do so, we first label automatically the PVs and LAA of every mesh with the help of the LDDMM framework and the varifold representation of shapes (see [7]). We compute the *Atlas*, or mean shape of the population, giving us a diffeomorphic registration from the Atlas to every shape. The population can then be fully represented by the Atlas T and a set of deformations $\{\varphi_i\}_{i \leqslant n}$. After hand-labeling of the Atlas, we warped the labels through the deformations in order to label every patient.

In practice, due to different types of vein anatomy in the population, the atria were separated in three classes to prevent big deformations from moving the labels too far from the roots of the veins. The analysis was performed with the `deformetrica` software [2] and the deformation and varifold kernels widths were both set to 10 mm.
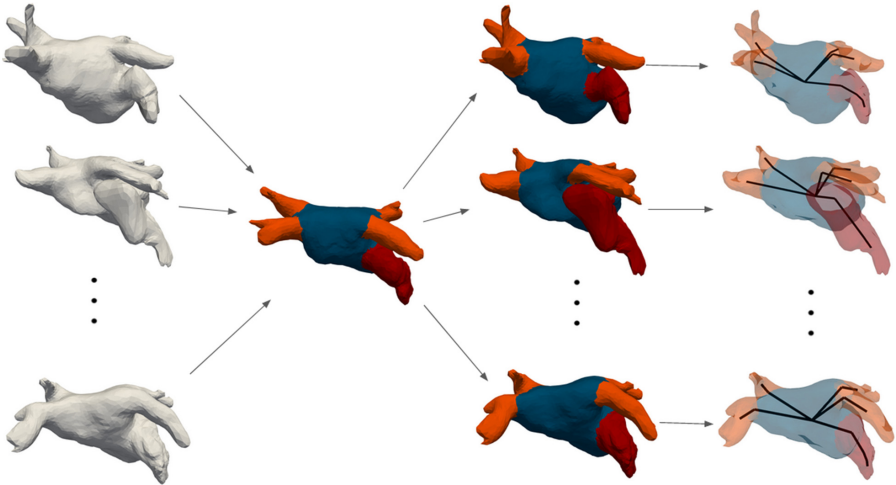

Fig. 1. Pipeline for the graph representation of the LA.

As a result, the variations of each mesh were captured well up to the residual noise on the surface. Therefore, atlas labels were warped to each subject faithfully to the anatomy of the atrium, as shown in Fig. 1.

---

[1] https://www.mmgtools.org/.

To achieve the graph representation, we extract the centre of mass of each label (i.e. PVs, LAA and body of LA) as well as the centre of each junction between labels, representing the ostium[2]. These points are the graph nodes, each branch is connected to the centre of the body of the LA.

Finally, to have a unified coordinate system between graphs, we set the body centre as the origin, fix the x-axis as the direction from the centre to the LAA ostium, and chose the left anterior PV ostium as the second direction.

2.3   Design of Fusion and Classification Loss Function

Let's denote $D = \{D_k\}_{k=1}^N$ N datasets with respective dimension d_k, and $(x, y) = \{x_k, y_k\}_{k=1}^N$ a set of pairs of observation x_k and label y_k from every dataset. Let $z \in \mathbb{R}^d$ be the latent variable shared by all elements of (x, y), with d the dimension of the latent space such that $d \ll \inf\{d_k | 1 \leqslant k \leqslant N\}$. We aim at having a common representation across datasets D_k, thus for every k a common distribution $p(z | x_k, y_k)$.

To do so, we use variational inference by introducing $\phi = \{\phi_k\}_{k=1}^N$ the inference parameters for each datasets, θ the common generative parameters and density functions $q_\phi(z | x_k, y_k) \in \mathcal{Q}$, which we want, on average, to be as close as possible to the common posterior $p_\theta(z | D)$. By using the Kullback-Leibler divergence, this problem translates to:

$$\underset{q \in \mathcal{Q}}{\operatorname{argmin}} \, \mathbb{E}_N[\mathcal{D}_{KL}(q_\phi(z | x_k, y_k) || p_\theta(z | D))] \tag{1}$$

Because of the intractability of $p_\theta(z | D)$ we cannot directly solve this optimisation problem. We aim to find a lower bound of (1) by expanding the Kullback-Leibler divergence:

$$\mathcal{D}_{KL}[q_\phi(z | x_k, y_k) || p_\theta(z | D))] = \int_{\mathbb{R}^d} q_\phi(z | x_k, y_k)[\ln q_\phi(z | x_k, y_k) - \ln p_\theta(z | D)]dz \tag{2}$$

Using Bayes' theorem, we can now rearrange the divergence to:

$$\mathcal{D}_{KL}[q_\phi(z | x_k, y_k) || p_\theta(z | D))] = \mathcal{D}_{KL}[q_\phi(z | x_k, y_k) || p_\theta(z | x))] \\ - \mathbb{E}_{z \sim q_{\phi_k}}[\ln p_\theta(y | z, x)] + \ln p_\theta(y | x) \tag{3}$$

Which yields the following evidence lower bound:

$$\ln p_\theta(y | x) - \mathcal{D}_{KL}[q_\phi(z | x_k, y_k) || p_\theta(z | D))] = \mathbb{E}_{z \sim q_{\phi_k}(z | x_k, y_k)}[\ln p_\theta(y | z, x)] \\ - \mathcal{D}_{KL}[q_\phi(z | x_k, y_k) || p_\theta(z | x))] \tag{4}$$

We impose this constraint over all datasets, by supposing that every dataset is conditionally independent, we have the following evidence lower bound:

$$\mathcal{L}(x, y, \theta, \phi) = \frac{1}{N} \sum_{k=1}^N \mathbb{E}_{z \sim q_{\phi_k}(z | x_k, y_k)} \left[\sum_{k=1}^N \ln p_\theta(y_k | z, x_k) \right] \\ - \mathcal{D}_{KL}[q_{\phi_k}(z | x_k, y_k) || p_\theta(z | x))] \tag{5}$$

---

[2] The ostium is the centre of the root of the vein.

Maximising the lower bound $\mathcal{L}$ is therefore equivalent to optimising the initial problem (1). The distribution $p_\theta(y_k|z, x_k)$ of shared parameters θ is learned by a common decoder from the latent space, and acts on the labels y_k, in this sense the decoder is a classifier on the set of all labels in D. In addition the learned distribution p_θ is a meta-parameter that contains information from every datasets in D.

Unlike variational auto-encoders, the reconstruction objective in (5) is over the labels y, which transforms the traditional decoder into a classifier. Moreover having more than one encoder impacts the reconstruction loss which becomes a cross reconstruction of the labels from every dataset. This constraint forces the encoders to identify a common latent representation across all the datasets.

In practice, we assume $\mathcal{Q}$ is the Gaussian family, parameters θ and ϕ are initialised randomly, the optimisation is done by stochastic gradient descent, using an adaptive learning rate with an Adam optimiser and back-propagation.

3 Synthetic Experiments

The aim of our synthetic experiments[3] is to highlight the possibilities offered by our method, for the sake of interpretability, we chose to use a relatively simple parameterisation for our model, consisting of a neural network with three fully connected layers as our encoder and 2 fully connected layers for the classifier. All activation functions are ReLU, and a Softmax function is used for classification.

The synthetic data was generated with the `make-classification` function from the `scikit-learn`[4] library, by generating clusters of points for multi-label classification by sampling from a normal distribution. We then generate similar datasets by applying transformation and adding noise to the initial problem. Thanks to this method, we can generate a high variety of independent datasets where the features have different distributions, but sharing the same target space.

We generate a collection of 5 independent datasets of dimensions varying between 10 and 20, sharing three labels. The latent dimension is set to 4 and we train our model jointly across datasets, as well as on each dataset separately. On 100 experiments, results show a generally increased overall accuracy score on test data when the datasets are trained jointly. With a 90% confidence level, the accuracy for joint training lies in $[0.83, 0.99]$ with a median of 0.92, while the accuracy for separate training lies in $[0.73, 0.94]$ with a median of 0.86, and about the same amount of epochs are required for convergence. Moreover, we can see in Fig. 2 that training multiple datasets jointly leads to a coherent latent space shared between labels (Fig. 2b), while trying to independently represent the latent spaces of each dataset in a common space leads to poor consistency of the latent space representation across datasets (Fig. 2a).

---

[3] The code is available at https://github.com/Inria-Asclepios/mcvc.

[4] https://scikit-learn.org/.

(a) The 5 independent latent spaces   (b) Our model's Joint latent space
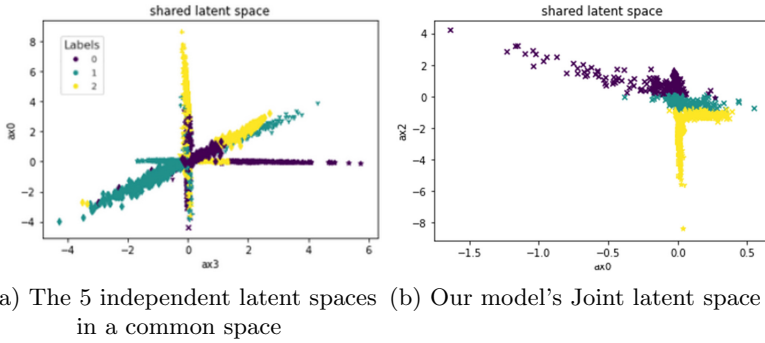   in a common space

Fig. 2. Latent space after convergence on 5 datasets with 3 labels. Our model provides a unified representation (Fig. 2b), not achievable when datasets are trained independently (Fig. 2a)

3.1   Noisy Labels

Because of the shared latent space between the collection of datasets D, some subspace is allocated for labels that may not belong to a given dataset D_i. This brings robustness to the classification when there is an uncertainty on the ground truth labels. In fact, if we assign a wrong label to a given dataset D_i the model is capable of assigning the observation to another label due to the constraint of obtaining a coherent latent space. To illustrate this point we created four datasets sharing three common possible labels; for one dataset we modified the third label into a fourth label. In Fig. 3 we show the evolution of the joint latent space throughout training, which highlights the preference for obtaining a coherent latent space rather than high-accuracy, the test accuracy dropped from 0.88 to 0.75 while the fourth label completely disappeared in the predictions.
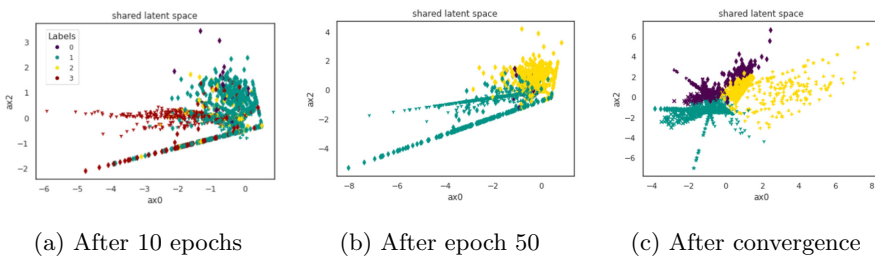

(a) After 10 epochs   (b) After epoch 50   (c) After convergence

Fig. 3. Evolution of joint latent space whith noisy labels. The wrong label (red) disappears completely from predictions to become coherent label (green) (Color figure online)

4   Application to LAA Graphs

Our clinical database is composed of 50 patients in systole (of which 27 are Thrombus positive) and 57 patients in diastole (of which 37 are Thrombus positive), thus we use two encoders in which we feed the nodes of the graphs to enable a joint analysis of the dataset. The labels are 1 if the patient is Thrombus positive and 0 otherwise. As such, the classes are well balanced for both datasets. We train the model with the same architecture and hyper-parameters as for synthetic data.

After 10-fold cross-validation, the model yields an average test accuracy of 0.89, for the diastolic set 0.92, and for the systolic set 0.86. In contrast, attempts to classify the subsets independently suffer from mode collapse; even a careful hyper-parameters fine-tuning results in very poor accuracy scores, reaching at best 0.65 in both cases. This highlights the robustness of the model and the clear advantage of joint analysis. Figure 4b shows the shared latent space with the systole subset as ($\times$) and diastole subset as ($\bullet$). We see a clear separation of classes as well as a common separator for the subsets, while systole and diastole are well grouped together for each class.

When we attempt to classify the complete dataset without any splitting, disregarding the considerable changes in shape during the cardiac cycle, we obtain slightly worse accuracy results of 0.84. In addition, such a model is less interpretable as important features for a given class can be contradictory in between subsets. To highlight this we investigate important features and show possible bio-markers.

As an additional baseline we performed PCA followed by logistic regression with cross validation and grid search on the number of principal components. We observed a much lower accuracy both on the complete dataset or individual subsets (0.65 at diastole, 0.71 at systole and 0.66 on the complete dataset).

We compare results from three interpretation algorithm (Integrated Gradient [11], DeepLIFT [10], and KernelSHAP [6]) implemented in `captum`[5] python library. We first feed the tests 100 times during cross validation and compute the feature importance algorithms on samples that are predicted right with more than a 95% certainty. Mean values over all samples are the final feature attribution scores. Figure 4a highlights the necessity to split the datasets to keep clinical coherence; It shows the Integrated gradient score on the x value of the Right Inferior PV (RIPV) point when the dataset is split (systole and diastole) and when it is trained commonly (i.e. not split). We see the model without splitting the data disregards this feature when in fact it is important for the diastole subset.

Figure 5 shows the possible bio-markers, in black are atlases of the population at systole and diastole, in blue are important feature values for control patients, in red are the ones for Thrombus positive patients. At diastole, the model seems to focus on the left PVs; their ostia being closer to and rotated towards the interior compared to the rest of the frame in Thrombus positive
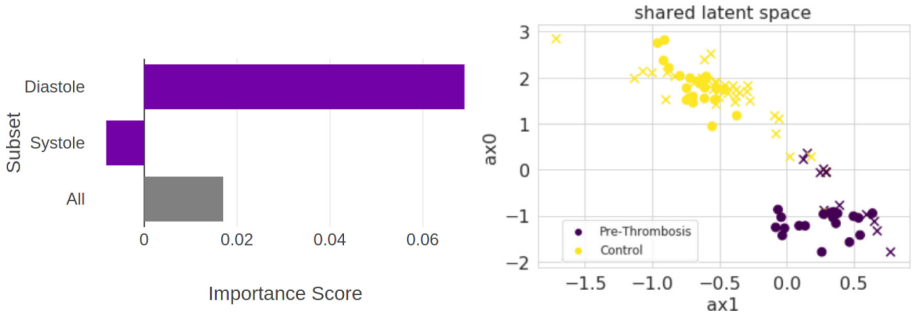

(a) Integrated Gradient scores of the RSPV center's x-axis for joint (purple) and without split (gray) training.

(b) Shared latent space.

Fig. 4. Results of the method. Joint analysis enables better interpretation of biomarkers (4a) as well as successful representation in the latent space (4b).

cases; In addition, for those cases, all veins intersection with the LA body are more horizontal, this 'fold' could impact the hemodynamic environment of the LA, which plays an important part in clot formations. At systole, the left PV and the angulation of the appendage are the focus. For Thrombus positive cases, PVs are rotated towards the interior with the left interior PV being on top and aligned with the LAA; The LAA ostium tends to be closer to the center and the LAA lower. Finally we see again the PVs tendency to being more 'folded'. While being interpretable, this also highlights the importance of analysing systole and diastole images separately.
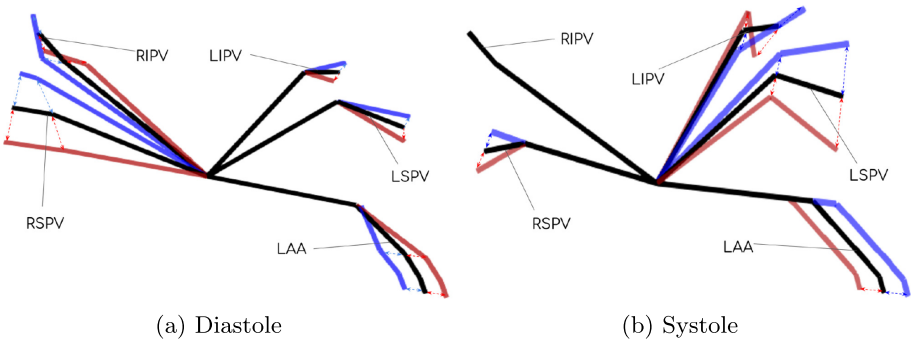
(a) Diastole   (b) Systole

Fig. 5. Visualisation of important features for predicting presence (red) and absence (blue) of thrombosis, both at diastole and systole. (Color figure online)

5   Conclusion

In this work, we provided a graph representation of the LA to analyse possible image-based bio-markers. In order to enable joint analysis of systole and diastole graphs we presented a new method at the crossroad between multi-task learning and meta learning to tackle the joint analysis of multiple heterogeneous datasets. By leveraging on the idea that the whole is better than its parts, we proposed a classification scheme with good interpretation properties of the latent space highlighted in the study of LA Graphs. We believe that the coherent latent space inherited from our model makes it possible to have deep neural network as encoders while conserving the interpretability of simpler models. We aim at further exploiting this property by applying the method to joint analysis of datasets containing much more heterogeneity. Finally, we believe the lightweight graph representation can be added in a more complete and multi-disciplinary study of the LA.

References

1. Antelmi, L., Ayache, N., Robert, P., Lorenzi, M.: Sparse multi-channel variational autoencoder for the joint analysis of heterogeneous data. In: International Conference on Machine Learning, pp. 302–311. PMLR (2019)
2. Bône, A., Louis, M., Martin, B., Durrleman, S.: Deformetrica 4: an open-source software for statistical shape analysis. In: Reuter, M., Wachinger, C., Lombaert, H., Paniagua, B., Lüthi, M., Egger, B. (eds.) ShapeMI 2018. LNCS, vol. 11167, pp. 3–13. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-04747-4_1
3. Di Biase, L., et al.: Does the left atrial appendage morphology correlate with the risk of stroke in patients with atrial fibrillation? Results from a multicenter study. J. Am. Coll. Cardiol. **60**(6), 531–538 (2012)
4. Hospedales, T., Antoniou, A., Micaelli, P., Storkey, A.: Meta-learning in neural networks: a survey. arXiv e-prints pp. arXiv-2004 (2020)
5. Jia, S., et al.: Automatically segmenting the left atrium from cardiac images using successive 3D U-Nets and a contour loss. In: Pop, M., et al. (eds.) STACOM 2018. LNCS, vol. 11395, pp. 221–229. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-12029-0_24
6. Lundberg, S., Lee, S.I.: A unified approach to interpreting model predictions. Adv. Neural. Inf. Process. Syst. **30**, 4765–4774 (2017)
7. Pennec, X., Sommer, S., Fletcher, T.: Riemannian Geometric Statistics in Medical Image Analysis. Academic Press, Amsterdam (2019)
8. Quinn, G.R., Severdija, O.N., Chang, Y., Singer, D.E.: Wide variation in reported rates of stroke across cohorts of patients with atrial fibrillation. Circulation **135**(3), 208–219 (2017)

9. Ruder, S.: An overview of multi-task learning in deep neural networks. arXiv e-prints pp. arXiv-1706 (2017)
10. Shrikumar, A., Greenside, P., Kundaje, A.: Learning important features through propagating activation differences. In: International Conference on Machine Learning, pp. 3145–3153. PMLR (2017)
11. Sundararajan, M., Taly, A., Yan, Q.: Axiomatic attribution for deep networks. In: International Conference on Machine Learning, pp. 3319–3328. PMLR (2017)
12. Watson, T., Shantsila, E., Lip, G.Y.: Mechanisms of thrombogenesis in atrial fibrillation: Virchow's triad revisited. Lancet **373**(9658), 155–166 (2009)
13. Yaghi, S., et al.: The left atrial appendage morphology is associated with embolic stroke subtypes using a simple classification system: a proof of concept study. J. Cardiovasc. Comput. Tomogr. **14**(1), 27–33 (2020)