2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Cardiac Disease Representation Conditioned by Spatio-temporal Priors in Cine-MRI Sequences Using Generative Embedding Vectors

Henry Peña, Santiago Gómez, David Romo-Bucheli and Fabio Martinez[1]

Abstract— **Cardiac cine-MRI is one of the most important diagnostic tools for characterizing heart-related pathologies. This imaging technique allows clinicians to assess the morphology and physiology of the heart during the cardiac cycle. Nonetheless, the analysis on cardiac cine-MRI is highly dependent on the observer expertise and a high inter-reader variability is frequently observed. Alternatively, the ejection fraction, a quantitative heart dynamic measure, is used to identify potential cardiac diseases. Unfortunately, this type of measurement is insufficient to distinguish among different cardiac pathologies. This quantification does not exploit all the heart functional information conveyed by cine-MRI sequences. Automatic image analysis might help to identify visual patterns associated with cardiac diseases in the cine-MRI sequences and highlight potential biomarkers. This paper introduces a conditional generative adversarial network that learns a mapping between the latent space and a generated cine-MRI data distribution involving information from five different cardiac pathologies. This net is guided from the left ventricle segmentation and the velocity field that is computed as *prior* information to focus on the deep representation of salient cardiac patterns. Once the deep neural networks are trained, a set of validation cine-MRI slices is represented in the embedding space. The associated embedding descriptor, in the latent space, is found by minimizing a reconstruction error in the generator output. We evaluated the obtained embedded representation as a disease marker by using different classification models in 16000 pathological cine-MRI slices. The representation retrieved by using the best conditional generative model configuration was used on the classifier models yielding an average accuracy of 90.04% and an average F1-score of 89.97% in the classification task.**

Clinical relevance— **Construction of a topological embedding space, from generative representation, that fully exploits hidden relationships of cine-MRI and represent cardiac diseases.**

I. INTRODUCTION

Heart disease is the leading cause of death around the world, some studies estimate that it will cause more than 23 million deaths in 2030 [1]. Cardiac cine magnetic resonance imaging (cine-MRI) allows radiologists to assess the morphology and ventricular function of the patient's heart. Nonetheless, such analysis is based on the training and expertise of the radiologists, and therefore a high inter-reader variability is frequently observed. Besides, some quantitative metrics are based on the global cardiac index that loses valuable Spatio-temporal information captured in volumetric cine-MRI sequences.

Recently, the Computer-Assisted Decision Support Systems (CADS) in cardiac cine-MRI imaging are mainly dedicated to segment ventricles to compute hemodynamic metrics, such as heart rate variability (HRV) or ejection fraction (EF) [2], [3]. These approaches are however limited to use already known pathological patterns on cine-MRI without providing new information to support diagnosis decisions. Moreover, these cardiac measurements can be sensitive to cardiac disease variability. For instance, the calculation of these hemodynamic patterns is dependent on proper ventricle segmentation and such analysis is limited to the end-diastole (ED) and end-systole (ES) phases. In consequence, most of the Spatio-temporal information available in the cine-MRI sequences is not currently used by state-of-the-art CADS.

Other strategies have focused on automatic heart disease classification. For instance, Zhan et al. computed several statistics related to the shape and textural features, and then sequential forward feature selection was used to improve the classification process [4]. Cetin et al. trained deep learning models to segment regions of the heart on MRI scans throughout the cardiac cycle [5]. Nonetheless, these supervised strategies do not exploit the Spatio-temporal information available in the cine-MRI sequences. Generative adversarial networks (GANs) promise to find uncovered descriptors and potentially new biomarkers associated with structural and dynamic changes in particular pathologies. Despite this potential, in cine-MRI, such generative architecture has been mainly adopted in the artificial generation of cardiac sequences[6], [7]. For instance, in [8] progressive sequential causal GANs (PSCGAN) is proposed to simultaneously synthesize and LGE-equivalent image and the segment diagnosis-related tissues. Other works on GANs are used to avoid problems of anonymity and privacy of the data [9]. Realistic GAN-generated images are used instead of real images to avoid sharing clinical imaging data [10], [11].

This work presents a strategy based on a conditional GAN architecture, conditioned by cardiac prior information, to recover embedding patterns and to differentiate among five different cardiac pathologies. Prior information related to ventricle shape and local velocity fields is used as conditional input fed into the GAN to find a suitable representation associated with cardiac conditions. As a result, meaningful mappings and embeddings are learned to represent cine-MRI data distribution This representation is discriminative among the different cardiac conditions, as shown by low-dimensional projections of the embeddings space. Moreover, a quantitative evaluation of the embedding was also carried out via classical machine learning algorithms. The classifiers successfully discriminated among five different heart cardiac conditions using the embedding representation with an accuracy of 90.04% with the best configuration.

[1] Biomedical Imaging, Vision and Learning Laboratory (BivL^{2}ab). Universidad Industrial de Santander (UIS), COLOMBIA famarcar@saber.uis.edu.co

II. Proposed approach

Figure 1 summarizes a schematic diagram depicting the main composing blocks of the proposed approach and the proposed evaluation. First, a generative representation is trained conditioned by the shape and motion priors (subsection II.A). From such representation, an embedding coding vector to compactly describe observed conditions on cine-MRI slices is recovered (subsection II.C). The whole computed embedding are mapped into a topological space to cluster samples and classify vectors, according to heart conditions.

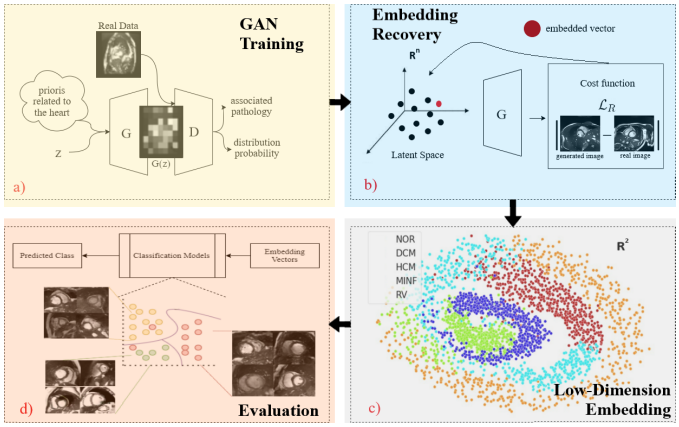

Fig. 1. Pipeline of the proposed method.(a) A conditioned GAN architecture learns a mapping between embeddings and generated/real cine-MRI images (b) Test cine-MRI images are represented by embedding vectors. (c) The embedding domain shows clear separation among five different cardiac conditions. (d) Evaluation via classifiers fed by the embedded representation

A. Deep conditioned and generative representation

This work uses the Multiple Conditional Input GAN (MCIGAN) architecture [12]. This network restricts the generative samples using multiple input conditions while avoiding artifact generation and mode collapse (see architecture illustration in Figure 2). As a typical GAN, this net is composed by two main deep neural networks: A generator and a discriminator. In the MCIGAN, the generator is fed with a random vector in the latent space, the associated heart pathology and vector information related with prior structural or motion information, embedded as a constraint. The discriminator is trained to distinguish if the input cine-MRI sequence was synthesized by the generator or if it is real. The discriminator outputs a probability of the cine-MRI sequence being real, and also a set of probabilities related to the different cardiac pathology classes.

The GAN training is carried out by using the Wasserstein cost function L_W, as:

$$L_W = E_{\tilde{x}\sim P_r}[D'(x)] - E_{\tilde{x}\sim P_g}[D'(\tilde{x})] \quad (1)$$

where $E_{x\sim P_r}$ is the expected value over all instances of the real data distribution P_r given the discriminator $D'(x)$. Contrary, the term $E_{\tilde{x}\sim P_g}$ counts probability to belong to generated images. An auxiliary loss function L_{aux} is then

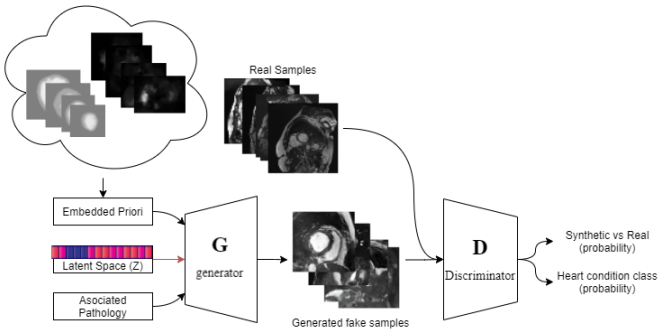

Fig. 2. Architecture of the generative adversarial network which has multiple inputs in order to evaluate the resulting distributions of different heart distributions.

introduced to include heart condition priors. In such case a typical measure of squared difference between the real and false pathologies c is fixed to obtain a "label class penalty". To further regularize the training process, a feature consistency lost L_{cons} is herein introduced by operating Laplacian matrices inputs dedicated to retain local details and preserve local structure of input images. Then a general minimization is used by integrating the whole defined losses, as:

$$\min_G L_W + L_{aux} + \lambda_{cons} L_{cons} \quad (2)$$

The structural and dynamic heart information is captured as prior information as follows:

1) Left ventricle shape: A U-Net was previously trained to segment the left ventricle in the ED and ES [13]. The motivation for using this masks is to encourage the learned representation in the GAN to focus on the heart morphology, which is a demonstrated biomarker of many cardiac conditions [14]. The segmentations were propagated throughout the entire cardiac cycle achieving a coherent and a dense representation of the temporal ventricle shape.

2) Motion flow patterns: The heart morpho-physiology is also a fundamental bio-marker for the characterization of abnormal or pathological conditions [15]. This information is directly related to the temporal deformation of the ventricles and their synchronized ability to pump specific volumes of blood. In this work, a dense optical flow algorithm was used at pixel scale [16]. The captured outgoing motions were associated with strong displacements between consecutive t and $t+1$ slices. This process results in a representation of the velocity of the cardiac chambers.

B. Embedding computation

Once the generative learned representation is fixed, computing the embedded vector for a new cine-MRI sequence consists on mapping it to the embedding domain. To obtain such embedding vector, an iterative process is carried out to minimize a residual loss function [17]. Initially, the method computes the loss between a target image x_n and an initial synthetic image $G(z_0)$ for a random vector z_0.

The loss is minimized via back-propagation, with a total of κ iteration steps, by providing the gradients to update the latent vector coefficients ($z_0 \rightarrow z_1 \rightarrow ... \rightarrow z_\kappa$). The embedding vector z_κ generates a synthetic image G_{z_κ} close enough to represent the image x_n. The loss function herein implemented is described as $\mathcal{L}_R(z_\gamma) = \sum |x_n - G(z_\gamma)|$ [18].

C. Data

The proposed methodology was validated over the public ACDC dataset [19]. A total of 100 cine-MRI volumes (around of 16000 slices), together with the associated segmentation were herein used. Apical slices (around 30% of each volume) were discarded in this study because the poor information related with cardiac information. The total of cine-MRI is classified into five cardiac conditions: myocardial infarction **(MINF)**, Dilated cardiomyopathy **(DCM)**, Hypertrophic cardiomyopathy **(HCM)**, Abnormal right ventricle **(RV)**, and control heart volumes **(NOR)**. For the statistical validation, the total set was split on 80% for the GAN training, and 20% for validation and latent space generation.

D. Experimental setup

The generative representation was set to receive and generate cine-MRI slices, scaled with spatial size of 64×64. The integrated input was configured as follows: 1) the embedding prior (shape and motion) coded as images of 64×64, 2) a random vector of 512 dimensions and 3) the pathology information encoded in a vector of three dimensions. In the training stage a total set of 12000 slices was included. Training mini-batches were set to 32. An early stopping criteria was used: if there is no difference in the training loss on consecutive iterations, the training is halted. Also, the maximum set of iteration was set to 1×10^7. The parameter λ_{cons} was fixed to 100. The rest of hyper-parameters were established following the same configuration that MCIGAN.

A topological low-dimensional embedding space was obtained with the UMAP algorithm [20]. The algorithm was applied on the validation set using the hamming distance. A quantitative validation was also carried out. Regarding the capability to separate among different cardiac diseases, three machine learning algorithms were used on the original embedding representation. The classification strategies here in selected were the K-Nearest Neighbors (KNN) (with $K = 13$), the Random Forest (RF) Classifier (115 trees), and the Support Vector Machine (SVM) (with RBF kernel). The whole models were adjusted with a Bayesian optimization process.

III. Evaluation and Results

The capability of the proposed strategy to encode embedding vectors representing the cardiac conditions was validated on the low-dimensional space, as well as in the classification task. The resultant embedding space was explored by using a set of embedding vectors generated from different priors. The selected embeddings were projected into a low dimensional space, using the UMAP strategy. Three different projections were obtained by using only the shape (segmentation), only the local movement (optical flow), and both the shape and motion prior. Figure 3 summarizes the three topological representation which properly separate the embedding vectors according to the different priors. The embedding space built from only prior motion information has some overlapping among the cardiac conditions. In contrast the projection using both priors shows clear separation among the cardiac conditions, which points to a discriminative representation of the cine-MRI slices.

A second experiment was carried out to quantitatively validate the capability of obtained vectors to describe heart pathologies. The original embedding vectors (with size of 512) were used to train machine learning classifiers and perform a classification task to separate the cardiac conditions. As done in the previous experiment, the classification was also carried out following three different prior configurations. Table I summarizes the results of three configurations following different accuracy metrics. It should be noted that in general the generative representation could be useful as biomarkers to differentiate among cardiac pathologies. In average, the three representation achieved an average precision of 0.8644, sensitivity of 0.8622 and a F1-score of 0.8633. Remarkably, the best performance was achieved by the generative model that include both: shape and motion prior information. Hence, embedding vectors seem to encode morphological and physiological information of particular cine-MRI sequences. Also, although there exist some differences between the classifiers performance, the results are quite similar.

TABLE I

CLASSIFICATION PERFORMANCE USING K-NEAREST NEIGHBOURS(KNN), RANDOM FOREST(RF) AND SUPPORT VECTOR MACHINE(SVM). THE NEXT METRICS WERE USED: PREC: PRECISION, SENS: SENSITIVITY, SPEC: SPECIFICITY, F1-SC: F1-SCORE.

	Heart Condition & Segm				Heart Condition & Flux				Heart Condition & Segm Flux			
	Prec	Sens	Spec	F1-Sc	Prec	Sens	Spec	F1-Sc	Prec	Sens	Spec	F1-Sc
KNN	0.89 ±0.05	*0.89 ±0.03*	*0.97 ±0.01*	0.89 ±0.03	0.83 ±0.05	0.83 ±0.01	0.95 ±0.01	0.83 ±0.06	*0.90 ±0.03*	*0.89 ±0.03*	*0.97 ±0.01*	**0.9 ±0.01**
RF	0.86 ±0.06	0.86 ±0.03	0.96 ±0.01	0.86 ±0.04	0.83 ±0.04	0.83 ±0.08	0.95 ±0.01	0.83 ±0.05	*0.88 ±0.04*	*0.88 ±0.02*	*0.97 ±0.01*	**0.88 ±0.01**
SVM	0.88 ±0.06	0.88 ±0.04	*0.97 ±0.01*	0.88 ±0.04	0.82 ±0.05	0.81 ±0.11	0.95 ±0.01	0.81 ±0.06	*0.89 ±0.02*	*0.89 ±0.02*	0.97 ±0.00	**0.89 ±0.01**

Finally, a more detailed analysis of the per-condition classification of the best generative representation, *i.e.,* including shape and motion information, is also shown on Figure 4. A confusion matrix was computed for the three different classifiers. As expected, the embedding vectors that represents cine-MRI sequences defined as normal, achieve the best classification rate w.r.t the other cardiac conditions. However, the other classes have a good performance, except for the DCM pathology. This fact may be attributed to the limitation of the generative model to reproduce dilated cardiac conditions.

IV. Concluding remarks

This paper presents a generative strategy applied to cine-MRI sequences that aim at a better cardiac representation by

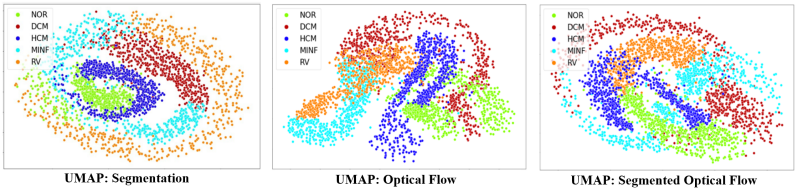

Fig. 3. Unsupervised UMAP clustering for the different priori information.

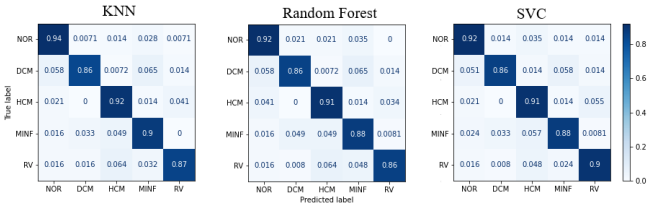

Fig. 4. Classification results were measured in accuracy using the MCIGAN-generated embedding space conditioned with Spatio-temporal priors. The confusion matrix of the (A) K-Nearest Neighbours, (B) Random Forest, and (C) Support vector machine models are shown above.

using a set of Spatio-temporal priors. The resulting hidden vectors could be used to support cardiac analysis in cine-MRI sequences. The results show a remarkable low-dimensional embedded space capable of accurately representing different cardiac conditions. A set of experiments using shape and motion priors shows that the integration of both information yields a better representation of cardiac conditions as shown by the F1-score metric results in table I. The proposed embedding vectors represented complex visual patterns in cine-MRI slices and obtained remarkable results when applied in a cardiac condition classification task.It should be reminded that pathologies used in this work are limited to known diseases in the medical domain. Another limitation in this work is that the quality of the apriori information is dependant on computational techniques used for extracting them. Future work should cover an important study on the explainability of the embedding space and the development of progressive architectures for the GAN, allowing a simpler training process than the current one,and finally, more focus should be placed on learning and studying the embedded space rather than learning the map between the distributions.

V. Acknowledgments

The authors acknowledge the Vicerrectoriá de Investigación y Extensión (VIE) of the Universidad Industrial de Santander for supporting this research work by the project: "Predicción de patologías cardíacas utilizando representaciones de aprendizaje profundo en secuencias de resonancia magnética cardíaca (CMR), with SIVIE code 2703.

References

[1] Mendis *et al.*, *Global atlas on cardiovascular disease prevention and control*. World Health Organization, 2011.

[2] E. Reinertsen *et al.*, "Heart rate-based window segmentation improves accuracy of classifying posttraumatic stress disorder using heart rate variability measures," *Physiological Measurement*, vol. 38, pp. 1061–1076, may 2017.

[3] L. Liang, W. Mao, and W. Sun, "A feasibility study of deep learning for predicting hemodynamics of human thoracic aorta," *Journal of Biomechanics*, vol. 99, p. 109544, 2020.

[4] Cetin *et al.*, "A radiomics approach to computer-aided diagnosis with cardiac cine-mri," in *Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges*, (Cham), pp. 82–90, Springer International Publishing, 2018.

[5] Zhang *et al.*, "Deep learning for diagnosis of chronic myocardial infarction on nonenhanced cardiac cine mri," *Radiology*, vol. 291, no. 3, pp. 606–617, 2019. PMID: 31038407.

[6] Wong *et al.*, "Understanding data augmentation for classification: when to warp?," in *2016 international conference on digital image computing: techniques and applications (DICTA)*, pp. 1–6, IEEE, 2016.

[7] L. Perez and J. Wang, "The effectiveness of data augmentation in image classification using deep learning," *arXiv preprint arXiv:1712.04621*, 2017.

[8] C. Xu *et al.*, "Contrast agent-free synthesis and segmentation of ischemic heart disease images using progressive sequential causal gans," *Medical Image Analysis*, vol. 62, p. 101668, 2020.

[9] Diller *et al.*, "Utility of deep learning networks for the generation of artificial cardiac magnetic resonance images in congenital heart disease," *BMC Medical Imaging*, vol. 20, p. 113, Oct 2020.

[10] G. Litjens *et al.*, "State-of-the-art deep learning in cardiovascular image analysis," *JACC: Cardiovascular Imaging*, vol. 12, no. 8_Part_1, pp. 1549–1565, 2019.

[11] Carneiro *et al.*, *Review of Deep Learning Methods in Mammography, Cardiovascular, and Microscopy Image Analysis*, pp. 11–32. Cham: Springer International Publishing, 2017.

[12] G. Yildirim, C. Seward, and U. Bergmann, "Disentangling multiple conditional inputs in gans," 2018.

[13] Ronneberger *et al.*, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, (Cham), pp. 234–241, Springer International Publishing, 2015.

[14] Beerbaum *et al.*, "Coronary anomalies assessed by whole-heart isotropic 3d magnetic resonance imaging for cardiac morphology in congenital heart disease," *Journal of Magnetic Resonance Imaging*, vol. 29, no. 2, pp. 320–327, 2009.

[15] D. Goksel, M. Ozkan, and C. Ozturk, "Cardiac motion analysis in mri for classification," in *Proceedings IEEE International Symposium on Biomedical Imaging*, pp. 935–938, 2002.

[16] T. Brox, C. Bregler, and J. Malik, "Large displacement optical flow," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 41–48, 2009.

[17] Radford *et al.*, "Unsupervised representation learning with deep convolutional generative adversarial networks," *arXiv preprint arXiv:1511.06434*, 2015.

[18] Schlegl *et al.*, "Unsupervised anomaly detection with generative adversarial networks to guide marker discovery," in *Proc. of IPMI*, pp. 146–157, Springer, 2017.

[19] Bernard *et al.*, "Deep learning techniques for automatic mri cardiac multi-structures segmentation and diagnosis: is the problem solved?," *IEEE transactions on medical imaging*, vol. 37, no. 11, pp. 2514–2525, 2018.

[20] L. McInnes, J. Healy, and J. Melville, "Umap: Uniform manifold approximation and projection for dimension reduction," 2020.