ISSN 0120-4157

Biomédica

Revista del Instituto Nacional de Salud

Citación provisional:

Moreno A, Bautista LX, Martínez F. Cardiac disease discrimination from 3d-convolutional kinematic patterns on cine-MRI sequences. Biomédica. 2024;44 (Supl. 1).

Recibido: 04-08-23

Aceptado: 13-02-24

Publicación en línea: 13-02-24

Cardiac disease discrimination from 3d-convolutional kinematic patterns on cine-MRI sequences.

Discriminación de enfermedades cardiacas utilizando patrones cinemáticos codificados con convoluciones 3d en secuencias de cine-RM

Clasificación de cardiopatías con descriptores cinemáticos

Alejandra Moreno, Lola Xiomara Bautista, Fabio Martínez

Biomedical Imaging, Vision, and Learning Laboratory (BIVL^{2}ab), Universidad Industrial de Santander, Bucaramanga, Colombia

Corresponding author:

Fabio Martínez Carrillo, Biomedical Imaging, Vision, and Learning Laboratory (BIVL^{2}ab), Universidad Industrial de Santander, Cl. 9 #27, Bucaramanga, Colombia

famarcar@saber.uis.edu.co

Author contributions:

Alejandra Moreno and Fabio Martinez conceived the method.

Alejandra Moreno implemented the method.

All authors conceived the evaluation experiments, analyzed results and writing the paper.

Introduction. Cine-MRI sequences are a key diagnostic tool for observing anatomical information, allowing experts to localize and determine suspicious pathologies. Nonetheless, such analysis remains subjective and prone to diagnosis errors.

Objective. To develop a binary and multi-class classification considering various cardiac conditions through a spatio-temporal model that highlights kinematic movements to characterize each disease.

Materials and methods. This research focuses on using a 3D convolutional representation to characterize cardiac kinematic patterns, during the cardiac cycle, which may be associated with pathologies. The kinematic maps are obtained from the apparent velocity maps computed from a dense optical flow strategy. Then, a 3D convolutional scheme learns to differentiate pathologies from kinematic maps.

Results. The proposed strategy was validated with respect to the capability to discriminate among myocardial infarction, dilated cardiomyopathy, hypertrophic cardiomyopathy, abnormal right ventricle, and normal cardiac sequences. The proposed method achieves 78.00% average accuracy and 75.55% average F1-score, respectively. Likewise, the approach achieved 92.31% of accuracy for binary classification between pathologies and control cases.

Conclusion. The proposed method is able to support the identification of kinematic abnormal patterns, associated with a pathological condition. The resultant descriptor, learned from the 3D convolutional net, preserves detailed spatio-temporal correlations and could emerge as possible digital biomarkers of cardiac diseases.

Keywords: heart diseases; diagnostic imaging; magnetic resonance spectroscopy.

Introducción. Las secuencias de cine-RM son una herramienta diagnóstica clave para observar la información anatómica, lo que permite a los expertos localizar y determinar aquellas patologías que resulten sospechosas. No obstante, este análisis sigue siendo subjetivo y propenso a errores de diagnóstico.

Objetivo. Desarrollar una clasificación binaria y multi-clase considerando diferentes condiciones cardiacas a través de un modelo espacio-temporal que permite resaltar los movimientos cinemáticos logrando caracterizar cada enfermedad.

Materiales y métodos. Este estudio se centra en el uso de una representación convolucional 3D para caracterizar los patrones cinemáticos durante el ciclo cardiaco, que pueden estar asociados a patologías. Para ello, se obtienen mapas cinemáticos a partir de los mapas de velocidad aparente calculados mediante una estrategia de flujo óptico denso. A continuación, un esquema convolucional 3D aprende a diferenciar patologías a partir de mapas cinemáticos.

Resultados. La estrategia propuesta se validó con respecto a la capacidad de discriminar entre pacientes con infarto de miocardio, miocardiopatía dilatada, miocardiopatía hipertrófica, ventrículo derecho anormal y pacientes normales. El método propuesto alcanza una precisión media del 78,00% y una puntuación F1 score del 75,55%, respectivamente. Asimismo, el enfoque alcanzó un 92,31% de precisión para la clasificación binaria entre patologías y casos de control.

Conclusiones. El método propuesto es capaz de apoyar la identificación de patrones cinemáticos anormales, asociados a una condición patológica. El descriptor resultante, aprendido de la red convolucional 3D, conserva

correlaciones espacio-temporales detalladas y podría surgir como posible biomarcador digital de enfermedades cardiacas.

Palabras clave: cardiopatías; diagnóstico por imagen; espectroscopía de resonancia magnética.

Associated cardiac pathologies represent the main cause of death worldwide, representing around 30% of the total deaths (1). The movement and kinematic components of cardiac structures represent a key biomarker of heart disorders. Magnetic resonance imaging (MRI) has become the primary clinical diagnostic technique for quantifying, inspecting, and analyzing the heart. For instance, from MRI modality, the ejection fraction can be calculated to discriminate among several cardiac conditions. However, the estimation of such measurements is based on manual delineation, which can be subject to errors. In addition, cardiac measurements may be insufficient to characterize and differentiate the diverse cardiac behaviors which are often complex among different cardiac diseases. Computational methods have allowed modeling and quantifying the motion and shape of cardiac features, supporting tasks related to segmentation (2-5), motion analysis, and classification of cardiovascular diseases. Regarding segmentation, the approaches have used atlas templates (2,6), encoder-decoder architectures (7-9), and even deep representations dedicated to localizing regions (3). Likewise, Qin et al. (10) uses a Motion-Seg Net to simultaneously obtain motion and shape estimations, under an unsupervised scheme. Additionally, semi-supervised learning was introduced to propagate cardiac disease labels, using as a backbone a U-Net that codifies the shape and motion features (11). Additionally, Punithakumar et. al. (12) calculates diverse statistics related to velocities and ventricle distances to classify pathologies, such as infarcts, dilated heart disease, and other cardiovascular diseases. Also, Zhen et. al. (13) performed an unsupervised cardiac image representation, learned from a multi-scale deep network, that achieved a direct volume estimation.

The present work introduces a deep volumetric representation which fully characterizes cardiac motion patterns, allowing to obtain motion embedding descriptors that classify diverse cardiac diseases. From deep cardiac representation, the high-level net embedding vectors are obtained as hidden kinematic cardiac descriptors used to classify and discriminate among several cardiac pathologies. In the next sections will be fully described the methodological approximations, as well as, the validation over a public dataset.

Materials and methods

This work introduces a 3D convolutional representation to encode cardiac kinematic maps as embedding descriptors with the capability to classify a set of cardiac conditions. From velocity fields, cardiac kinematic maps are calculated to locally represent patterns such as normal acceleration, tangential acceleration, divergence, and vorticity. These enriched and dense motion primitives are convolved several times to obtain a hierarchical deep representation of (2D+t) spatio-temporal patterns through the cardiac cycle along the short axis. The main hypothesis that underlies this work is the capability of spatio-temporal motion patterns to represent particular cardiac conditions. In consequence, the architecture is able to receive ($2D + t$) feature maps of the whole cardiac cycle generating a hidden deep and latent representation that discriminates among different cardiac diseases. The general pipeline of the proposed approach is shown in figure 1.

Kinematic cardiovascular maps

To characterize the motion patterns in the cardiac cycle, it is necessary to build a set of kinematic maps that recover motion features from a dense optical flow

strategy. The displacement vector field, computed among consecutive frames, is related to the apparent velocity of the cardiac cycle. Here, it was selected as an approximation of the optical flow that recovers large displacements, as well as, a deformation representation that lies in a constraint across nearby regions. For each couple of consecutive frames $I(x)_t$, $I(x)_{t+1}$ a dense motion field was computed $\vartheta \coloneqq (u, v)^T$. From this motion field, we computed $x = (x, y)^T$ a respective displacement vector (u, v) for each pixel. Hence the dense motion field is obtained as a typical minimization of appearance $(\|I_t - I_{t+1}\|^2)$ and gradient $(\|\nabla I_t - \nabla I_{t+1}\|^2)$ with a function that matches non-local points (SIFT points) where the flow region is coherent (14). The kinematic maps are then derived from this optical flow field and are represented as $K = [k_1, k_2, \dots k_i]$. In this case, K represents a motion feature map, while i is the index of each calculated kinematic (velocities, accelerations, divergences, or vorticities). It is important to note that these feature maps can be used as isolated observations or even integrated to enrich cardiac disease representation from motion patterns.

Initially, we consider two acceleration types, where the first one is the normal acceleration (NA), which represents the direction change of the velocity taking as reference the local center of rotation of the point analyzed $\left(a_N(t) = \frac{\|v(t)\|}{\|\vartheta_T'(t)\|}\right)$. Also, the tangential acceleration (TA) was introduced to approximate heart deformation during the cardiac cycle, as: $(a_T(t) = \frac{u}{dt}\|v(t)\|)$. In addition,the divergence (DIV) was calculated, which measures the motion density of the input compared to the output. The divergence is mathematically formulated as: $(div(t) = \frac{\partial u(t)}{\partial x} + \frac{\partial v(t)}{\partial y})$. Also, the vorticity (VOR) was included in this study, which measures the cardiac

rotational motion during contraction (ED) and relaxation phases (ES) defined as: $(vor(t) = \frac{\partial v(t)}{\partial x} - \frac{\partial u(t)}{\partial y})$.

A deep 3D convolution architecture

A key issue in cardiac conditions analysis is the modeling and quantification of spatial and temporal patterns, which allows to stand out correlations between sequence observations and specific pathologies. In this work, a robust 3D convolutional representation is implemented, which captures spatio-temporal patterns at different processing levels, through a hierarchical convolutional configuration (15). Hypothetically, we assume that a cine-MRI observation can be fully expressed by spatio-temporal patterns. These patterns are learned into a deep representation, which in turn is adjusted through a conditional discrimination rule. Interestingly enough, the proposed architecture is able to receive kinematic maps, allowing us to code more complex relationships related to cardiac conditions.

A sequence of images Φ of dimension $(L\ x\ H\ x\ W)$ represents either a cine-MRI sequence, the corresponding kinematics representation (velocities, accelerations, vorticity or divergence) or even a concatenation of multiple kinematics. In this case, L denotes the temporal frame number belonging to the cardiac cycle, and (H, W) the spatial frame dimensions. This sequence is then used as input in the convolutional representation and operated at different 3D convolutional layers. In such case, κ represents the (z, v, w)-dimensional convolution kernel, where the z dimension convolves over the temporal axis and the (v, w) dimensions over the spatial axes. At each processing step, we calculated a representation volume D'

which represents a bank of spatio-temporal feature maps, capturing a more accurate characterization of motion throughout the cycle (figure 2).

The total feature volume represents the union of each map b for the specified number D' of kernels. The respective generalization at different layers allows obtaining a multi-scale motion representation providing a signature for each cardiovascular disease.

A cardiac embedding representation

The proposed convolutional representation has the capability to predict cardiac conditions under an end-to-end scheme, fixing in the last layer a probability prediction output, according to training cardiac classes. Besides, the final layers of the architecture model are a set of hidden and complex relationships of kinematic inputs, representing a descriptor of a particular disease. In this work, we explore the embedding space that results from these descriptors and measures the capability to discriminate among pathologies. Specifically, these embedded vectors were extracted from the last dense connected layer. For a particular dataset, the inputs are then mapped to a trained net and the compact vectors X^i are recovered, with the label corresponding to the disease y^i. The set of training samples was used in order to create a random forest, defined as: (X, Θ_j), where Θ_j represents each decision tree of independent and identically distributed random variables being formed by a uniform random selection of characteristics. In fact, a particular threshold (τ_i) is learned for each kinematic (K), which create a node in the tree $\varrho(K, \tau_i)$ and build a new partition in the feature space. The group of trees gives an independent vote for predicting pathology, allowing drawing a discrete partition

over topological space to define the boundaries of cardiac classes and obtain an automatic classification.

Experimental setup

This section details the data and methods of the proposed approach to validate our method and its performance according to the classification task. The next subsection describes each of the components in the experimental setup.

ACDC database: This work was trained and validated with a public dataset named "Automated Cardiac Diagnosis Challenge" (ACDC) (16). The dataset consists of a set of cine-MRI images from patients diagnosed with cardiovascular diseases and a control population. Four pathologies are characterized by ejection fraction (EF) and other morphological features. The myocardial infarction (MINF) is defined by multiple myocardial segments with an abnormal contraction and a left ventricular EF of less than 40%. The dilated cardiomyopathy (DCM) is characterized as having a left ventricular EF of less than 40% and a diastolic left ventricular volume greater than 100 mL/m2. Several myocardial segments with a thickness greater than 15 mm in diastole, a left ventricular cardiac mass greater than 110 g/m2, and a normal EF constitute an indicator of hypertrophic cardiomyopathy (HCM). On the other hand, when a patient has a right ventricular cavity volume greater than 110 mL/m2 and a right ventricle EF lower than 40%, it indicates an abnormal right ventricle (RV) cardiac condition. Also, the dataset includes patients labeled with normal cardiac condition (NOR).

For whole image sequences recorded in the dataset, the heart position is mainly on the basal and the mid-cavity. Each patient has a mean of 9 slices (from apical to basal), varying from [13-56] temporal frames across the cardiac cycle. The study

includes 100 patients diagnosed with one of the described pathologies (20 patients for each cardiac condition). From such volumes of the dataset, it was considered a total of 1300 slices. From a data analysis study, the cardiac cycle for each volume was fixed in 13 temporal frames. Volumes with larger cardiac cycles were sub-sampled, ensuring to cover the end diastole and end systole.

Implementation details: The configuration of the proposed 3D convolutional architecture is summarized in table 1. All these convolutions have the same size ($2x2x2$), except to the first one ($1x2x2$). For the introduced method, two different strategies of classification were configured, described as follows:

- End-to-end-training: In this configuration, the 3D convolutional approach was fixed with a softmax layer to carry out the classification of cardiac pathologies. In this case, the net was trained with a batch of one, a learning rate of 0.001, and an Adam optimization. In this case, the proposed net was also adjusted with a dropout of 0.4, and batch normalization to prevent over-fitting and regularize the loss. We use 20 epochs at each run and follow a binary classification rule.
- Random forest classifier: In this strategy, we used the activations from embedding layer of learned net. We expect that such embedding encodes learning kinematic features, and allows discrimination among cardiac conditions. Particularly, we take the embedding descriptor from the layer $Dense_1$, which in turn serves as input to a random forest. In such cases, all the kinematics were trained independently, for each discrimination rule, between two cardiac conditions. Hence, each image was mapped to each architecture, obtaining the respective embedding vector, which corresponds

to the last layer. The corresponding embeddings are concatenated, representing the new cardiac descriptor of each input sample. A fine-tuning was performed with the parameters: 1) A maximum tree of 100 and a maximum depth of 60. 2) Each tree is encoded in a binary classification for experiments to discriminate between pair of classes, and 3) each tree encodes a multi-class classification between normal cardiac sequences vs. any cardiac pathology.

Statistical validation

Statistical validation: the proposed strategy was validated according to a “leave-one-patient-out” scheme. This scheme was adopted from the classical leave-one-out cross-validation. Particularly, in this scheme validation, one patient (with 13 slices) is left out for testing purposes. In contrast, the rest of the patients (in our particular experiment, 39 patients accounting for 507 slices for each binary classification) are used for training until all patients are validated. For end-to-end experiments, at each fold is trained a convolutional net, which after that is validated with samples of a particular patient. The average of results corresponds to reported performance on classification.

When the validation scheme is finished, a prediction for each patient is retrieved, helping to then account for each metric classification such as accuracy, precision, sensitivity, or F1-Score. It is important to note that at each iteration there was no overlap between patients, indeed we used only one cardiac cycle per patient. The input samples have a dimension of $(12, 128, 128, 1)$. For multiple kinematics experiments, the multidimensional input was set as $(12, 128, 128, 3)$. Each dimension corresponds to the cardiac cycle, height, width, and the concatenation

between kinematics. The best multi-kinematic representation was considered to obtain descriptor vectors to carry out experiments from embedding representation.

Results

Classification from an end-to-end scheme

Figure 3 illustrates corresponding heart kinematic activations from two different layers of the proposed architecture. As expected, these maps enhance spatial relationships that eventually may correspond to patterns associated with a specific disease. The illustrated sample corresponds to cine-MRI labeled as a MINF condition. As input, there were included independently the optical flow channels and the divergence. Also, the normal acceleration, divergence, and vorticity combination were mapped to the trained architecture to obtain deep hierarchical activations. For each illustrated input, these activations achieve consistent localization that stands out in particular kinematic behaviors at ventricles during a cardiac cycle. These activations hierarchically code cardiac descriptors that allow supporting an automatic classification but also, they can be implemented as observational maps to further analysis during diagnosis and clinical routine. For three input configurations, the activations of the first layer stand out local cardiac patterns while the $(L-1)$ layer focus on the coarse characterization of heart regions. These maps involve temporal correlations allowing an enriched description of the heart during the cardiac cycle.

Table 2 summarizes the classification performance of the proposed approach using independent kinematic cine-MRI features. The accuracy and the F1-score were the metrics selected to globally analyze the performance of the proposed descriptor. As observed, each independent kinematic has the capability to

discriminate between two cardiac conditions, being potential descriptors to support disease diagnosis. On average, the velocity field patterns (accuracy of 75.83% and F1-score of 71.50%) and the divergence (accuracy of 75.23% and F1-score of 72.34%) achieve better discrimination for whole experiments. These findings may be associated with principal components of heart dynamics, such as the rotation movements to describe the left ventricle and particular spatial velocity patterns, along the cardiac cycle. Interestingly enough, each kinematic excels in discrimination between a couple of conditions. For instance, the vorticity has remarkable results to classify between MINF vs DCM, while the TA and NA kinematics have a notable performance to separate DCM from control samples.

In a subsequent experiment, the most promising kinematic features were combined as an input block. Table 3 summarizes the results of different kinematic configurations. Considering the correlated nature of such kinematics (differential relationships from optical flow field), there is not a significant enhancement in global accuracy. Nonetheless, there exist some remarkable configurations, such as the DCM vs NOR that achieve an average accuracy of 92.50% and an F1-score of 92.68%, using the coupled configuration of the kinematics: TA, DIV, VOR.

Classification of embeddings from Random Forest

An additional multi-modal kinematic configuration was considered to better exploit the deep representation of each motion feature map. In this experiment, the new cardiac descriptor was evaluated with a random forest classifier. Table 4 shows an experiment using the late fusion of embedding vectors taken from deep representations of the next kinematics: normal acceleration, divergence, and vorticity. Following this configuration, the best accuracy result achieves an average

of 78.00% and an F1-score of 77.55%. Also, we should highlight that some experiments achieved a perfect classification score, showing the discrimination capability of the three deep kinematic representations.

The proposed representation can be also implemented as a triage alternative to classify between cine-MRI with any condition or control sequences. To validate such an alternative, the proposed approach was validated in an experiment that merged together all labels that correspond to the abnormal cardiac conditions in the same class (MINF, DCM, HCM, and RV). In such case, we obtain a binary classification from abnormal conditions vs control sequences, as seen in Table 5. The group embedding vectors that correspond to TA, DIV, and VOR kinematics were employed to carry out a late fusion classification. In this experiment, the proposed approach achieved a remarkable 92.31% and 91.19% of accuracy and F1-score, respectively.

Discussion

The proposed approach introduced a novel 3D convolutional net to quantify and characterize diverse spatio-temporal motion patterns on the complete cardiac functional cycle. This strategy is able to recover kinematic maps and obtain a hierarchical deep multi-level representation, built from a discrimination rule between cardiac conditions. Furthermore, we validated the classification and characterization capabilities of the 3D network operated on this cardiac kinematics. Complementary, we tested compact embedding outputs, which were thereafter used to train and validate a random forest classifier, achieving remarkable results. In fact, the estimation of the deep kinematic representation improved accuracy by over 6% and F1-Score by 10.88% for each kinematic with respect to the original

cine-MRI sequences. From this proposed kinematic setup, it is possible to combine and enrich motion representation by convolving multiple kinematics, at the same time, for a particular 3D net. This enriched representation (from normal acceleration, divergence, and vorticity) proves a better performance to discriminate among multiple cardiac conditions.

A remark in this approach is the capability to build compact embedding descriptors that code cardiac conditions and form a topological space, from which, it is possible to access an automatic classification. These resultant embedding correspondences may emerge as potential digital biomarkers of cardiac conditions, storing complex correlations, achieved from a learning optimization. This approach is promising to implement in a clinical routine to support triage protocols because of the exhibited performances of around 92.31% to discriminate between control and any cardiac condition, included in this study. Although CMRI is not currently the primary diagnostic study for most cardiac conditions, its growing advantages are becoming increasingly evident, leading to its adoption as a triage scheme for detecting specific cardiac conditions (17). In fact, current reports also evidence an effort to introduce such artificial intelligence tools in clinical protocols (18). Aditionally, the kinematic maps and resultant activations at different layers of hierarchical representation may be important during observational analysis.

In the state of the art, much of the methodologies are dedicated to perform ventricle segmentation tasks. From such resultant volumes are computed classical indexes, such as the ejection fraction, and the ventricle volume, among others (4,5,8,9). These indexes are computed from relative differences between end-of-diastole and end-of-systole. For instance, Puyol et. al. proposed a multi-modal

atlas that integrates MRI and ultrasound (US) to extract Laplacian motion descriptors, allowing the classification of patients with dilated cardiomyopathy from control subjects (6). Also, Yang et. Al proposed a registration strategy to quantify displacement of LV among temporal consecutive images, but bypassing RV analysis that might be remarkable for some diseases (7). Also, Clough et al. recovered variational embeddings to discriminate among cardiac diseases (17). Despite the remarkable contributions of these approaches, they remain dependent on proper ventricle segmentation to characterize cardiac pathologies. A main issue of these schemes is the dependency on guided segmentation and the loss of temporal patterns that may be crucial to enrich diagnosis. Likewise, these descriptors are based on known physical features, poorly exploiting potential hidden relationships that may be computed from the $(2D + t)$ information, available in complete cine-MRI sequences. In contrast, the proposed approach exploits motion relationships that may be computed from kinematic representation maps but also learned through a 3D representation. This strategy recovers complex motion patterns and may be useful to complement typical indices to support expert characterizations of particular cardiac conditions. In this line, some approaches have also captured motion patterns from left ventricles but again depending on a proper geometry recovery (7,10).

These experiments evidence a potential use of this strategy as triage support of patients in clinical schemes. Regarding the state-of-the-art, the proposed approach evidence competitive results regarding accuracy and precision by using several kinematic characteristics, following a cross-validation leave-one-patient-out. This fact shows the robustness of the embedding representation, which allows a reliable

classification of cardiac conditions. Besides, this approach operates without any segmentation requirement, which results interesting to achieve a more generalized heart representation from cine-MRI sequences. Also, the proposed approach was validated over an open database with real cine-MRI sequences over five different cardiac conditions. Nonetheless, the capability of the proposed approach should be validated over larger cohorts of data, from different clinical centers. In such a sense, it is expected to report the generalization capacity and the impact of each of the kinematic maps, regarding the discrimination capability among cardiac conditions. Also, a major exploration of middle and end embeddings should be carried out, exploring alternative descriptors of heart observations. For instance, a topological analysis or a geometrical search over embedding space may be an alternative to validate the discrimination capability. The proposed approach also requires additional processing schemes to include multi-classification from an end-to-end scheme. Future works include the study of other types of kinematics that can help to extract relevant patterns, such as attention feature maps. Finally, validation with a larger dataset that includes expert cardiologist annotations and clinical information will be considered to define a possible correlation with medical findings and the advantages and limitations of the approach.

This work proposed a deep volumetric convolutional net to classify cardiac pathologies from MRI sequences. The proposed approach computes kinematic maps, which allow deep representation to encode complex and hidden kinematics related to the observed pathologies. In fact, two classification schemes were used to validate the proposed approach: 1) from an end-to-end scheme, and 2) using embedding descriptors which further are mapped to a random forest classifier. The

proposed approaches evidence coherent competitive results over an open-access dataset. Future works include the study of geometrical embedding space and the validation with larger data cohorts that allow to establish the statistical scope to discriminate among close cardiac pathologies.

Conflict of interest

The authors have no conflict of interests to declare that are relevant to the content of this article.

Funding

This research was supported receiving funding from the Vicerrectoría de Investigación y Extensión (VIE) of the Universidad Industrial de Santander under the project entitled “Predicción de patologías cardíacas utilizando representaciones de aprendizaje profundo en secuencias de resonancia magnética cardíaca (CMR)”, with the SIVIE code 2703.

References

1. World Health Organization. World Health Statistics 2021. Geneva: World Health Organization; 2021.

2. Cetin I, Sanroma G, Petersen SE, Napel S, Camara O, Ballester MA, et al. A radiomics approach to computer-aided diagnosis with cardiac cine-MRI. arXiv. 2019:1909.11854v1. https://doi.org/10.48550/arXiv.1909.11854

3. Chang Y, Jung C. Automatic cardiac MRI segmentation and permutation-invariant pathology classification using deep neural networks and point clouds. Neurocomputing. 2020;418:270-9. https://doi.org/10.1016/j.neucom.2020.08.030

4. Chang Y, Song B, Jung C, Huang L. Automatic segmentation and cardiopathy classification in cardiac MRI images based on deep neural networks. IEEE

International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2018;1020-4. https://doi.org/10.1109/ICASSP.2018.8461261

5. Khened M, Kollerathu VA, Krishnamurthi G. Fully convolutional multi-scale residual DenseNets for cardiac segmentation and automated cardiac diagnosis using ensemble of classifiers. Med Image Anal. 2019;51:21-45. https://doi.org/10.1016/j.media.2018.10.004

6. Puyol-Anton E, Ruijsink B, Gerber B, Amzulescu MS, Langet H, De Craene M, et al. Regional Multi-View Learning for Cardiac Motion Analysis: Application to Identification of Dilated Cardiomyopathy Patients. IEEE Trans Biomed Eng. 2019;66:956-66. https://doi.org/10.1109/TBME.2018.2865669

7. Yang D, Wu P, Tan C, Pohl KM, Axel L, Metaxas DN. 3D Motion Modeling and Reconstruction of Left Ventricle Wall in Cardiac MRI. Funct Imaging Model Heart. 2017:10263:481-92. https://doi.org/10.1007/978-3-319-59448-4_46

8. Isensee F, Jaeger PT, Full PM, Wolf I, Engelhardt S, Maier-Hein KH. Automatic cardiac disease assessment on cine-MRI via time-series segmentation and domain specific features. arXiv.2017:1707.00587v2. https://doi.org/10.48550/arXiv.1707.00587

9. Wolterink JM, Leiner T, Viergever MA, Išgum I. Automatic Segmentation and Disease Classification Using Cardiac Cine MR Images. arXiv. 2017:1708.01141v1. https://doi.org/10.48550/arXiv.1708.01141

10. Qin C, Bai W, Schlemper J, Petersen SE, Piechnik SK, Neubauer S, et al. Joint motion estimation and segmentation from undersampled cardiac MR image. arXiv. 2019:1908.07623v1. https://doi.org/10.48550/arXiv.1908.07623

11. Zhen X, Wang Z, Islam A, Bhaduri M, Chan I, Li S. Multi-scale deep networks and regression forests for direct bi-ventricular volume estimation. Medical Image Analysis. 2016;30:120–9. https://doi.org/10.1016/j.media.2015.07.003

12. Kumaradevan Punithakumar, Ismail Ben Ayed, Islam A, Aashish Goela, Ross I, Chong J, et al. Regional heart motion abnormality detection: An information theoretic approach. Medical Image Analysis. 2013;17:311–24. https://doi.org/10.1016/j.media.2012.11.007

13. Zheng Q, Delingette H, Ayache N. Explainable cardiac pathology classification on cine MRI with motion characterization by semi-supervised learning of apparent flow. Medical Image Analysis. 2019;56:80–95. https://doi.org/10.1016/j.media.2019.06.001

14. Brox T, Malik J. Large Displacement Optical Flow: Descriptor Matching in Variational Motion Estimation. IEEE Trans Pattern Anal Mach Intell. 2011;33:500-13. https://doi.org/10.1109/TPAMI.2010.143

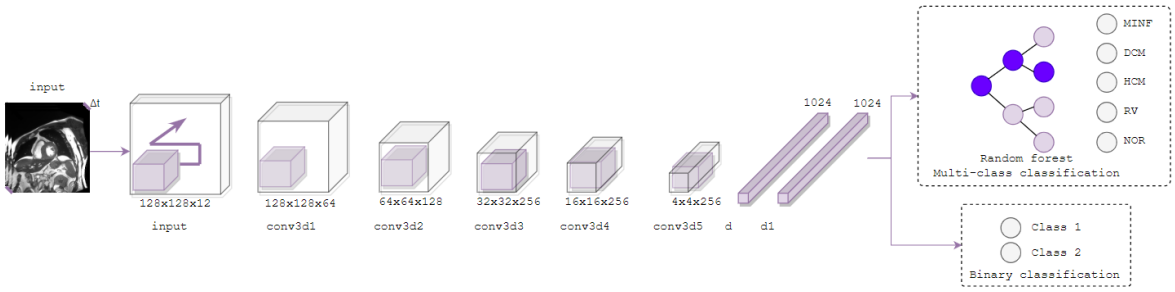
15. Gül Varol, Laptev I, Schmid C. Long-term temporal convolutions for action recognition. IEEE Trans Pattern Anal Mach Intell. 2018;40:1510-1517. https://doi.org/10.1109/TPAMI.2017.2712608

16. Bernard O, Lalande A, Zotti C, Cervenansky F, Yang X, Heng PA, et al. Deep Learning Techniques for Automatic MRI Cardiac Multi-Structures Segmentation and Diagnosis: Is the Problem Solved? IEEE Trans Med Imaging. 2018;37:2514-25. https://doi.org/10.1109/TMI.2018.2837502

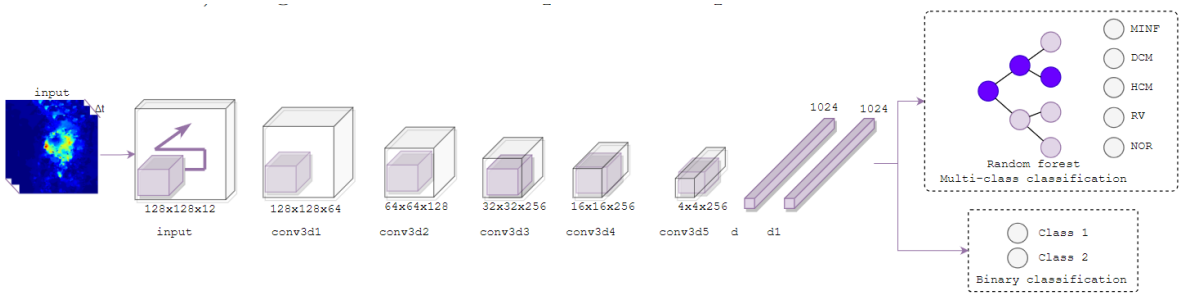
17. Broncano J, Bhalla S, Caro P, Hidalgo A, Vargas D, Williamson E, *et al*. Cardiac MRI in patients with acute chest pain. Radiographics. 2021;41:8-31. https://doi.org/10.1148/rg.2021200084

18. Fotaki A, Puyol-Antón E, Chiribiri A, Botnar R, Pushparajah K, Prieto C. Artificial intelligence in cardiac MRI: is clinical adoption forthcoming? Front Cardiovasc Med. 2022:8:818765. https://doi.org/10.3389/fcvm.2021.818765

19. Clough JR, Oksuz I, Puyol-Antón E, Bram Ruijsink, King AP, Schnabel JA. Global and Local Interpretability for Cardiac MRI Classification. Springer. 2019; 1767 LNCS:656–64. https://doi.org/10.1007/978-3-030-32251-9_72



A) Original cine-MRI sequence as input of the network



B) Normal acceleration cine-MRI sequence as input of the network.

Figure 1. Pipeline of the proposed representation to classify heart conditions from cine-MRI temporal sequences (bottom-up scheme) or using kinematic maps as input on deep representation.

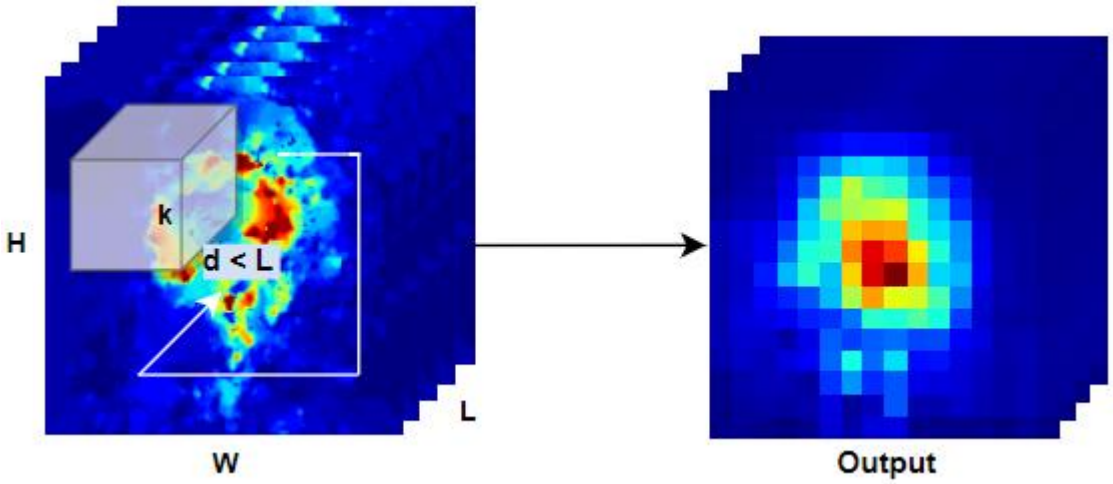


Figure 2. 3D convolution representation. An input image of size $L\ x\ H\ x\ W$ performing a convolution with a kernel $t * y * x$

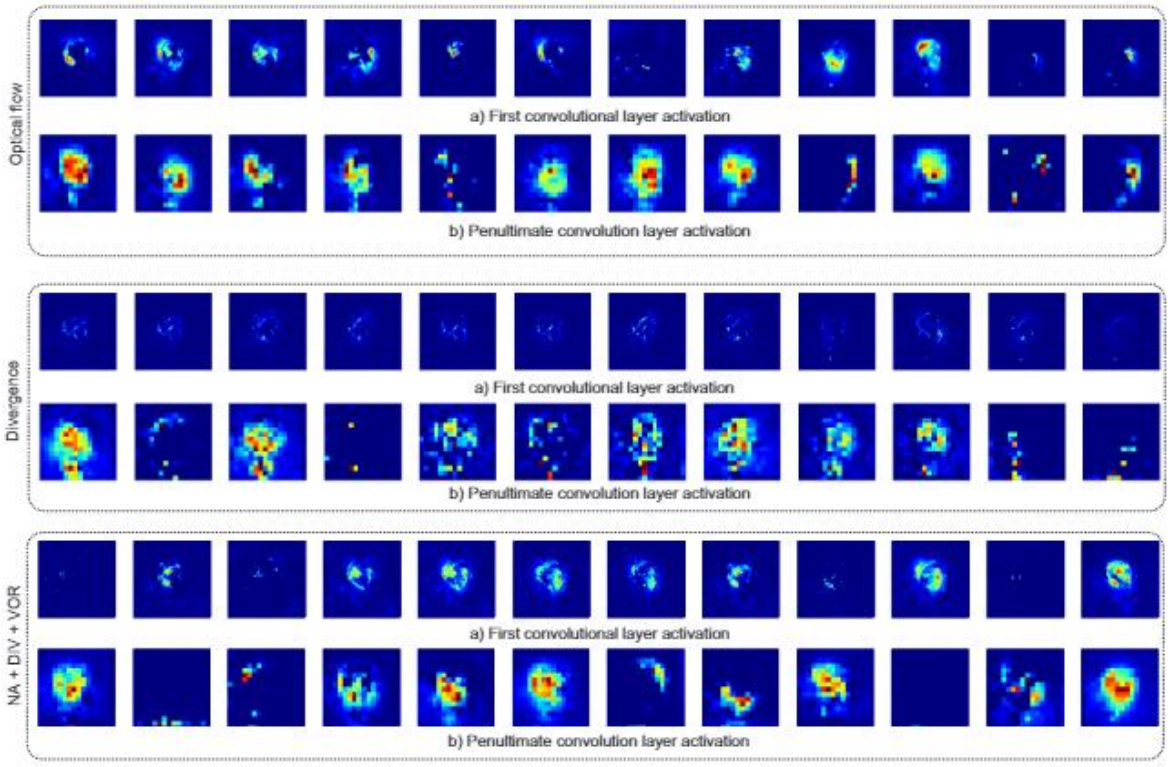


Figure 3. Feature map representation in the convolutional layers obtained from both the first and penultimate layers. These primitives are: the optical flow, divergence, and a concatenation between normal acceleration, divergence, and vorticity. The illustrated sample corresponds to cine-MRI labeled as a myocardial infarction.

Table 1. Parameters of the 3D deep convolutional architecture

Layers	Output shape	Parameters	Activation
$Input$	(12,128,128,1)	-	ReLU
$Conv3D$	(12, 128, 128, 64)	5842	ReLU
$Conv3D_1$	(6, 64, 64, 128)	221312	ReLU
$Conv3D_2$	(3, 32, 32, 256)	884992	ReLU
$Conv3D_3$	(2, 16, 16, 256)	1769728	ReLU
$Conv3D_4$	(1, 4, 4, 256)	1769728	ReLU
$Dense$	1024	1049600	ReLU
$Dense_1$	1024	1049600	ReLU
$Dense_2$	2	2050	ReLU

Table 2. Accuracy and F1-score obtained using the ACDC dataset in the deep learning strategy. ACC accounts for Accuracy and F1 for F1-score. The cardiac conditions are: the dilated cardiomyopathy (DCM), hypertrophic cardiomyopathy (HCM), abnormal right ventricle (RV), the myocardial infarction (MINF), and normal conditions (N).

Cardiac diseases	$a_T(t)$		$a_N(t)$		Divergency		Vorticity		Optical flow		Cine-MRI	
	ACC	**F1**	**ACC**	**F1**	**ACC**	**F1**	**ACC**	**F1**	**ACC**	**F1**	**ACC**	**F1**
MINF vs DCM	82,50	83,72	80,00	80,00	70,83	77,57	**85,00**	**84,21**	**89,20**	**85,00**	81,20	72,50
MINF vs HCM	65,20	61,53	70,00	70,00	70,2	77,57	62,50	61,53	73,09	65,0	59,09	57,50
MINF vs RV	80,00	77,77	80,00	78,94	80,30	80,00	85,00	78,94	81,33	77,50	70,53	70,00
MINF vs N	72.50	70.27	72.50	71.72	**87.59**	**87.59**	77.50	74.28	72.05	70.00	76.71	70,00
DCM vs HCM	55.00	52.63	62.50	59.45	59.40	73.81	55.00	50.00	80.71	72.50	60.38	65.00
DCM vs RV	67.50	64.86	60.00	61.90	72.56	67.54	65.00	65.00	72.78	62.50	51.43	55.00
DCM vs N	**85.00**	**84.21**	**85.00**	**85.00**	80.30	60.99	85.00	85.00	73.38	72.50	**83.38**	**80.00**
HCM vs RV	62.50	63.41	70.00	68.42	84.67	74.00	62.50	61.53	78.38	75.00	31.95	50.00
HCM vs N	75.00	73.68	72.50	73.17	81.25	64.25	65.00	61.11	70.00	67.50	75.76	67.50
RV vs N	75.00	73.68	72.50	70.27	65.15	55.05	67.50	66.66	67.40	67.50	57.65	55.00
MEAN	71.75	70.58	72.50	71.89	**75.23**	**72.34**	70.50	68.83	**75.83**	**71.50**	64.81	57.95

Table 3. Accuracy and F1-score obtained by using the ACDC dataset in the deep learning strategy taking into account diverse concatenations among time. The kinematic maps herein considered are a_N: Normal acceleration, a_T: Tangential acceleration div: divergency, and vor: vorticity. The cardiac conditions are: the dilated cardiomyopathy (DCM), hypertrophic cardiomyopathy (HCM), abnormal right ventricle (RV), the myocardial infarction (MINF), and normal conditions (N)

	$(a_N(t), a_T(t), div(t))$		$(a_N(t), a_T(t), vor(t))$		$(a_N(t), div(t), vor(t))$		$(a_T(t), div(t), vor(t))$	
Cardiac disease	**ACC**	**F1**	**ACC**	**F1**	**ACC**	**F1**	**ACC**	**F1**
MINF vs DCM	80.00	80.00	77.50	79.06	82.50	82.05	80.00	80.02
MINF vs HCM	62.50	61.54	62.50	63.41	60.00	57.89	65.00	66.67
MINF vs RV	**85.00**	**85.00**	85.00	85.00	85.00	85.00	80.00	77.76
MINF vs N	75.00	76.19	85.00	83.34	77.50	75.67	80.00	78.94
DCM vs HCM	55.00	52.63	50.00	47.36	55.00	52.63	52.50	53.65
DCM vs RV	70.00	68.42	67.50	66.67	77.50	79.06	67.50	68.29
DCM vs N	**85.00**	**85.00**	**90.00**	**89.47**	**90.00**	**89.47**	**92.50**	**92.68**
HCM vs RV	67.50	66.67	62.50	65.11	67.50	69.76	55.00	59.09
HCM vs N	75.00	75.00	87.50	87.80	80.00	78.94	75.00	73.68
RV vs N	80.00	80.00	85.00	84.21	82.50	82.92	85.00	84.21
MEAN	73.50	73.05	75.25	75.14	75.75	75.34	73.25	73.50

Table 4. The accuracy obtained using the ACDC dataset in the binary embedding classification with random forest. a_N: Normal acceleration, div: divergency, vor: vorticity. The cardiac conditions are: the dilated cardiomyopathy (DCM), hypertrophic cardiomyopathy (HCM), abnormal right ventricle (RV), the myocardial infarction (MINF), and normal conditions (N).

	$(a_N(t), div(t), vor(t))$			
Cardiac disease	**Accuracy**	**F1-Score**	**Precision**	**Recall**
MINF vs DCM	**100.00**	**100.00**	**100.00**	**100.00**
MINF vs HCM	80.00	78.10	85.00	80.00
MINF vs RV	**100.00**	**100.00**	**100.00**	**100.00**
MINF vs N	80.00	81.90	90.00	80.00
DCM vs HCM	80.00	80.00	86.67	80.00
DCM vs RV	60.00	66.30	86.67	60.00
DCM vs N	60.00	60.00	60.00	60.00
HCM vs RV	60.00	60.00	60.00	60.00
HCM vs N	80.00	71.11	64.00	80.00
RV vs N	80.00	78.10	85.00	80.00
MEAN	78.00	77.55	**81.73**	78.00

Table 5. The accuracy obtained using the ACDC dataset in the multi-class embedding classification with random forest between normal cardiac sequences w.r.t any cardiac disease.

Cardiac diseases	Accuracy	F1-Score	Precision	Recall
Cetin et al. (2017) [2]	**94.00**	**-**	**94.00**	***93.00***
Insensee et al. (2017) [9]	92.00	-	92.00	92.00
Khened et al. (2017) [5]	90.00	**-**	83.40	**100.00**
Wolterink et al. (2017) [10]	86.00	-	84.00	91.00
Ours	**92.31**	91.19	***92.95***	92.31