

HABILITATION A DIRIGER DES RECHERCHES

présentée devant

l'Institut National des Sciences Appliqués de Lyon
et l'Université Claude Bernard Lyon I

Analyse d'images médicales pour la fiabilisation des mesures cliniques en imagerie cardiaque

Medical image analysis for the reliability of clinical measurements in cardiac imaging

par

Olivier BERNARD

Docteur en traitement du signal et de l'image

Soutenue le 06 Septembre 2019 devant la commission d'examen

Composition du jury

<i>Président :</i>	Pr. Mireille Garreau	Université de Rennes 1, France
<i>Rapporteurs :</i>	Pr. Isabelle Bloch	Telecom Paris, France
	Dr. Jocelyne Troccaz	CNRS, Grenoble, France
	Pr. Christian Jutten	Université Grenoble Alpes, France
<i>Examineurs :</i>	Pr. Hervé Liebgott	Université de Lyon 1, France
	Pr. Denis Friboulet	INSA de Lyon, France

Mis en page avec la classe thesul.

Acknowledgments

Firstly, I would like to express my gratitude to all the institutions that made this work possible, namely the French national research agency (ANR), the French Laboratory of excellence (LABEX), the European Research Council through the European Union's Seventh Framework Program and the China Scholarship Council (CSC) for funding my research projects and then also the INSA of Lyon for providing me with an exceptional teaching environment with talented and motivated students and CREATIS for all these years of intense but excellent research.

Secondly, to all the jury members, thank you for your willingness to devote time to this work by reading this manuscript and wishing to make it better through your comments and questions. I must especially thank Denis Friboulet and Hervé Liebgott for following my research during all these years and stimulating me with interesting opportunities and projects. Denis, you were my PhD supervisor and then my team leader. Your good humour has always made my work environment conducive to my professional development. For the countless meetings, scientific discussions and laughs, I would never be grateful enough. Hervé, what can I say! You're my brother in arms. I never thought you could spend so many intense moments with a colleague in a work environment. I can't help but think back to all those conferences we had together... all those schemes I had to put in place to get us home safe and sound, especially to Prague :) You have always been a great source of inspiration to me, both from a human and a scientific point of view.

Third, to all my colleagues, thank you all for all these long and fruitful scientific discussions, the coffee breaks (coffee that I finally replaced with tea! Damn it, who would have bet it one day...), conferences, winter, spring and summer schools and everything else. To Damien, my older spiritual brother, I'm so excited about the projects we're putting in place. For sure, my scientific center of gravity will never be far from you in the coming years. Thank you also for stimulating me to resume sport during the lunch break, my psychological balance is all the better for it. To Thomas, your complete mastery of computer science is extremely inspiring for me and a source of intellectual stimulation. To Sorina, your calm, wisdom and kindness are pillars on which I do not hesitate to rely, thank you! To all my PhD students, Thomas, Daniel, Miaomiao, Yitian, Joao, Sarah, Yunyun and Ewan, what adventures passed with you all! Be assured that my research would not have been as intense and tasty without you. I know that sometimes I may seem very demanding and picky, but know that I am very proud of what you have achieved during your thesis and what you have become.

Last but not least, to my family for their support over the years. To my parents who have always believed in me. To Carole, my wife and partner through every struggle and conquest, thank you so much for everything. You have followed me all these years with such patience and faith in me that I realize how lucky I am to have you at my side. I am blessed to receive your love and your example of love and care for all those around us is something I can only hope to achieve. I find that together with Laura and Alexis, we are a great team all four of us, you are my reason for living and I would be nothing without you.

To all of those above and the ones I have forgotten to mention, my greatest thanks!

Olivier Bernard

*To Laura, Alexis and Carole
my sun ray, my energy and my beloved*

Table of contents

Personal presentation	1
------------------------------	----------

Curriculum vitae	3
-------------------------	----------

1	Training	3
2	Position - Experience	4
3	Editorial responsibilities	4
4	Awards	4
5	Teaching activities	5
5.1	Summary table of my teaching activities	5
5.2	Teaching activities during my PhD (147 hours in total)	6
5.3	Teaching activities during my one year associate professor (ATER) position (196 hours in total)	7
5.4	Teaching activities as a full-time associate professor at INSA-Lyon (220 hours on average per year)	7
6	Collective responsibility	9
6.1	Teaching	9
6.2	Research	10
6.3	Associate professor recruitment jury	10

Supervision and scientific impacts	11
---	-----------

1	Supervision activities	11
1.1	PhD theses	11
1.2	Masters of research	21

2	Scientific impacts	22
2.1	International collaborations	22
2.2	National collaboration	22
2.3	Awards and distinctions	22
2.4	Associate editor positions	22
2.5	Reviewing work	23
2.6	Organization of international challenges	23
2.7	Regular involvement to scientific events	23
2.8	Publications and communications	24

Research activities	37
----------------------------	-----------

General introduction	39
-----------------------------	-----------

Cardiac segmentation : toward robust volumetric indices estimation	41
---	-----------

1	Introduction	41
2	State-of-the-art in cardiac image segmentation	42
2.1	Cardiac ultrasound	42
2.2	Cardiac MRI	44
2.3	Previous cardiac datasets	45
3	1 st contribution : setting up open-access datasets	46
3.1	CAMUS dataset	47
3.2	CETUS dataset	51
3.3	ACDC dataset	54
3.4	Reference segmentation and contouring protocol	55
4	2 nd contribution : B-spline Explicit Active Surface paradigm (BEAS)	58
4.1	Methodology	58
4.2	Automatic initialization strategy	61
4.3	Framework description	64
4.4	Experiments	66
4.5	Results	67

4.6	Discussion	71
5	3 rd contribution : deep convolutional neural network study on RT2DE	75
5.1	Motivations	75
5.2	Evaluated methods	75
5.3	Results	80
5.4	Discussion	96
5.5	Conclusions	98
6	4 th contribution : deep convolutional neural network study on cardiac MRI	100
6.1	Motivations	100
6.2	Evaluated methods	100
6.3	Solutions for automatic cardiac diagnosis	102
6.4	Results	102
6.5	Segmentation Challenge	103
6.6	Discussion	111
6.7	Clinical implications	115
6.8	Conclusions	116
7	Summary - Conclusions	117

Cardiac motion estimation : toward robust strain estimation	121
--	------------

1	Introduction	121
2	State-of-the-art in cardiac motion estimation	123
2.1	Cine MR	123
2.2	Tagged MR	124
2.3	Echocardiography	126
3	State-of-the-art in cardiac simulation	127
3.1	Cine MR image simulation	127
3.2	Tagged MR image simulation	128
3.3	Echocardiography image simulation	128
4	1 st contribution : realistic synthetic US and MRI sequences from the same virtual patients	131
4.1	Motivations	131
4.2	Overview	131
4.3	Template image sequences	133
4.4	MR modeling	133
4.5	Cardiac motion	136
4.6	Template Warping	136
4.7	US simulation	141

4.8	Result	143
4.9	Discussion	154
4.10	Conclusions	155
5	2 nd contribution : anatomical deformable model for strain estimation in cardiac imaging	156
5.1	Motivations	156
5.2	Anatomical mesh model	156
5.3	Tracking methodology	159
5.4	Results	162
5.5	Discussion	163
5.6	Conclusions	166
6	Summary - Conclusions	168
Ultrasound image reconstruction : toward new imaging techniques		171
1	Introduction	171
2	State-of-the-art in ultrafast image reconstruction	172
2.1	Principle of coherent compounding	174
2.2	Motion compensation methods for ultrafast imaging	175
3	Revisiting the influence of transmission scheme involved in ultrafast imaging with plane waves	177
3.1	Motivations	177
3.2	Methodology	178
3.3	Experiments	179
3.4	Conclusions	185
4	1 st contribution : Ultrasound Fourier Slice Beamforming	186
4.1	Motivations	186
4.2	Methodology	187
4.3	UFSB : compounding imaging from steered plane waves	189
4.4	UFSB : summary and practical implementation	190
4.5	Results	192
4.6	Numerical simulations	192
4.7	<i>In vitro</i> and <i>in vivo</i> experiments	194
4.8	Discussion	198
4.9	Conclusions	201
5	2 nd contribution : extension of Fourier-based techniques for ultrafast imaging with diverging waves	202
5.1	Motivations	202

5.2	Methodology	202
5.3	Experiments	206
5.4	Results	208
5.5	Discussion	215
5.6	Conclusions	217
6	Summary - Conclusions	218

Research perspectives	219
------------------------------	------------

French summary	225
-----------------------	------------

1	Résumé de mes travaux de recherche	226
2	Perspectives de recherche	230

Bibliography	235
---------------------	------------

List of figures

1	Typical images extracted from the proposed CAMUS database. Endocardium and epicardium of the left ventricle and left atrium wall are shown respectively in green, red and blue. [Left] input images ; [Right] corresponding manual annotations.	50
2	Variability in the quality of the volumes acquired in clinical conditions	52
3	Manual segmentation of the left ventricle at the level of the most basal slice. a) Original image. b) Image with the left ventricle (in blue) and the myocardium (in green).	56
4	Manual segmentation of the right ventricle at the level of the most basal slice. a) Original image. b) Right ventricle segmentation in red. Only the posterior part was considered because of the absence of contraction at the level of the anterior area (white arrow) between diastole and systole (it is thus not considered as a part of the ventricle).	56
5	Manual segmentation of the right ventricle (RV) at the level of the most basal slice. The left ventricle is in blue, the myocardium in green and the right ventricle in red. The first column displays original images while each row corresponds to a given patient. b) RV's border is drawn when there is no clear separation between the cavity and the aorta. d) No RV's border is drawn when there exists a clear separation between the cavity and the aorta.	57
6	Mean statistical shape models at ED (a) and ES (b).	60
7	Left ventricular long axis (LAX) and base detection : (a) Image pre-processing (red - edge indicator, green - circle center probability) ; (b) Maximum probability path, displayed over P_{max} ; (c) Base detection (red - sliding averaging plane ; green ball detected base position).	62
8	Ellipsoid fitting : (a) Feature map and optimal ellipsoid fit in cylindrical coordinates (green) ; (b and c) Initial surface for the 3-D image segmentation step. . .	62
9	Conceptual description of the proposed segmentation and tracking framework. First, automatic initialization is applied to the ED frame (A). The first stage of segmentation is then performed using BEAS (B). The result from this segmentation is used to detect the SAX orientation (C) and this information is then used to perform the second stage of segmentation using BEAS and the ED SSM. The final ED segmentation is then propagated frame to frame using the IAAOF (E) and a final refinement to the ES frame is performed using BEAS and the ES SSM (F).	65
10	Bullseye plots of average MAD and HD at every region for ED and ES for the training datasets. Measures in mm.	69

11	Best (a,c) and worst (b,d) automatic segmentation results (red) compared to manual contours by experts (green) at ED (top row) and ES (bottom row) from the CETUS training set. The three orthogonal planes shown for each 3D image were chosen according to the automatically defined LAx/SAx orientation.	69
12	Influence of the distance and angle error from the reference position and orientation on the distance metrics (MAD, HD and Dice) at ED. Horizontal dotted line indicates the performance obtained with the proposed automatic framework on the CETUS training set.	70
13	Influence of the variation of each of the considered parameters α , β and γ on the distance metrics (MAD, HD and Dice) at ED (blue) and ES (red dotted). Vertical dotted line indicates the preset parameter value. \star , \dagger and \ddagger indicate respectively that the difference to the result with the preset values was statistically significant at a $p < 0.05$, $p < 0.01$ and $p < 0.001$ level.	71
14	d_m and d_H distributions derived from the U-Net 1 and 2 results obtained on LV_{Endo} and LV_{Epi} structures.	83
15	Segmentation results obtained by the U-Net 1 architecture on Patient 27 (image defined as good quality). Ground-truth contours are dotted and prediction contours are drawn in full line.	84
16	Segmentation results obtained by the U-Net 2 architecture on Patient 27.	85
17	Segmentation results obtained by the SRF method on Patient 27.	86
18	Segmentation results obtained by the BEASM method with automatic initialization on Patient 27.	87
19	Segmentation results obtained by the BEASM method with semi-automatic initialization on Patient 27.	88
20	Segmentation results obtained by the cardiologist 2 on Patient 27.	89
21	Segmentation results obtained by the cardiologist 3 on Patient 27.	90
22	Tukey box plots computed from the geometrical results of the U-Net 1 architecture for three different schemes (GM for learning to simultaneously segment all three structures from good and medium image quality, $mono$ for learning to segment one structure from all image quality, $multi$ for learning to simultaneously segment all three structures from all image quality). Blue numbers correspond to mean values computed from each set of measurements. p-values are based on the Wilcoxon signed rank test computed with the $multi$ strategy as reference.	93
23	Box plots results computed from the U-Net 1 architecture for three different schemes (GM for the learning from good & medium image quality, $mono$ for the learning of one structure, $multi$ for the learning of the three structures at the same time). Blue figures displayed under each box correspond to mean values computed from each set of measurements. p-values are based on the Wilcoxon signed rank computed with the $multi$ strategy as reference.	94
24	Geometric scores of the three cardiologist-specific models on 400 patients (1600 images)	96

-
- 25 Automatic segmentation of a healthy subject (NOR) from the testing database. Each row corresponds to a given slice of the same volume at end diastole (ED). From top to bottom : basal slice toward apical slice. From left to right : input image, ground-truth, Isensee *et al.* (winner of the segmentation contest), Khened *et al.* (winner of the classification contest). For this patient (full volume at ED), Isensee *et al.* obtained a Dice score of 0.97, 0.92 and 0.96 and a Hausdorff distance of 2.2 mm, 3.1 mm, and 3.1 mm, respectively for the left ventricle, myocardium and right ventricle. Khened *et al.* got a Dice score of 0.97, 0.92 and 0.97 and a Hausdorff distance of 3.1 mm, 3.1 mm, and 3.1 mm, respectively for the left ventricle, myocardium and right ventricle. 105
- 26 Automatic segmentation of a patient with a myocardial infarction (MINF) with altered left ventricular ejection fraction from the testing database. Each row corresponds to a given slice of the same volume at end diastole. From top to bottom : basal slice toward apical slice. From left to right : input image, ground-truth, Isensee *et al.* (winner of the segmentation contest), Khened *et al.* (winner of the classification contest). For this patient (full volume at ED), Isensee *et al.* obtained a Dice score of 0.97, 0.91 and 0.97 and a Hausdorff distance of 2.7 mm, 3.9 mm, and 4.9 mm, respectively for the left ventricle, myocardium and right ventricle. Khened *et al.* got a Dice score of 0.97, 0.90 and 0.95 and a Hausdorff distance of 3.1 mm, 4.1 mm, and 10.7 mm, respectively for the left ventricle, myocardium and right ventricle. 106
- 27 Automatic segmentation of a patient with an hypertrophic cardiomyopathy (HCM) from the testing database. Each row corresponds to a given slice of the same volume at end diastole. From top to bottom : basal slice toward apical slice. From left to right : input image, ground-truth, Isensee *et al.* (winner of the segmentation contest), Khened *et al.* (winner of the classification contest). For this patient (full volume at ED), Isensee *et al.* obtained a Dice score of 0.92, 0.89 and 0.97 and a Hausdorff distance of 11.6 mm, 10.5 mm, and 5.9 mm, respectively for the left ventricle, myocardium and right ventricle. Khened *et al.* got a Dice score of 0.97, 0.92 and 0.96 and a Hausdorff distance of 4.6 mm, 10.0 mm, and 11.7 mm, respectively for the left ventricle, myocardium and right ventricle. . . . 107
- 28 Automatic segmentation of a patient with dilated cardiomyopathy (DCM) from the testing database. Each row corresponds to a given slice of the same volume at end diastole. From top to bottom : basal slice toward apical slice. From left to right : input image, ground-truth, Isensee *et al.* (winner of the segmentation contest), Khened *et al.* (winner of the classification contest). For this patient (full volume at ED), Isensee *et al.* obtained a Dice score of 0.98, 0.91 and 0.90 and a Hausdorff distance of 4.1 mm, 4.3 mm, and 10.7 mm, respectively for the left ventricle, myocardium and right ventricle. Khened *et al.* got a Dice score of 0.97, 0.87 and 0.92 and a Hausdorff distance of 10.4 mm, 11.6 mm, and 10.5 mm, respectively for the left ventricle, myocardium and right ventricle. 108

29	Automatic segmentation of a patient with abnormal right ventricle (ARV) from the testing database. Each row corresponds to a given slice of the same volume at end diastole. From top to bottom : basal slice toward apical slice. From left to right : input image, ground-truth, Isensee <i>et al.</i> (winner of the segmentation contest), Khened <i>et al.</i> (winner of the classification contest). For this patient (full volume at ED), Isensee <i>et al.</i> obtained a Dice score of 0.97, 0.89 and 0.97 and a Hausdorff distance of 24.3 mm, 24.1 mm, and 12.8 mm, respectively for the left ventricle, myocardium and right ventricle. Khened <i>et al.</i> got a Dice score of 0.97, 0.86 and 0.95 and a Hausdorff distance of 24.3 mm, 24.1 mm, and 12.9 mm, respectively for the left ventricle, myocardium and right ventricle.	109
30	Confusion matrix of the winner of the classification challenge [?].	111
31	Average Dice index and Hausdorff distances for every method reported in Table 13 broken down for every pathology.	114
32	Histogram of degenerated slices ED (left), and ES (right).	115
33	Typical degenerated result at the base of the heart. [Left] input image; [Middle] ground truth; [Right] prediction.	115
1	Synthetic echocardiographic images used in [?] (a), [?] (b) and [?] (c). All techniques present intrinsic limitations which limit their usability as benchmarking tools. The solution in (a) employs a simplistic speckle model (<i>i.e.</i> high speckle intensity inside the muscle and low intensity outside) leading to unrealistic binary-like images. In (b) the simulated cine-loop is obtained by warping the B-mode pixel images of a real recording, thus producing texture warping artifacts. In (c) the presence of warping artifacts is avoided by using an ultrasound simulator. Yet, the reference motion field is computed by applying an elastic registration technique to the real recording working as template and therefore <i>i)</i> no control on the simulated motion is possible, <i>ii)</i> errors of the registration algorithm are directly reflected on the simulated motion and, moreover, <i>iii)</i> the benchmark displacement is biased towards the elastic registration technique employed. The pipeline described in this paper overcomes such limitations	130
2	Proposed pipeline for the simulation of realistic cardiac US, cine and tagged MR sequences.	132
3	Illustration of the different steps involved to simulate the initial T_1 and T_2 maps relative to the first frame (ED).	134
4	Illustration of the use of the generic transformation model. The ED cine MR image was warped to ES by the generic transformation. The grid superimposed on the image shows the corresponding deformation. The myocardium and the background are shown in red and green respectively. The generic transformation was computed from the ED/ES simulation meshes and the static background points.	137

5	Illustration of the temporal and spatial alignments described in Section 4.6.2 and 4.6.3 respectively. The myocardium is segmented at ED. The static background points are shown as black dots surrounding the mesh. The segmentation mesh is first tracked in the template sequence, yielding the tracking meshes \mathcal{R}_t shown in red (the upper row). Then, the E/M model is exploited for simulating myocardial deformations starting from the same segmentation mesh, leading to another sequence of meshes \mathcal{S}_k shown in the lower row. Each simulation frame k is mapped to a cardiac time $\phi(k)$ of the template through the temporal alignment. The spatial alignment $\mathcal{T}_{\{\mathcal{R}_{\phi(k)}, \mathcal{B} \rightarrow \mathcal{S}_k, \mathcal{B}\}}$ then register the template image to the corresponding simulation mesh \mathcal{S}_k	139
6	The computation of the Gaussian shaped weighting function.	140
7	Simulated cine MR images. Three slices located at basal, mid and apical regions of the left ventricle are shown. The mesh contour is superposed on the images.	144
8	Simulated tagged MR images. Three channels with orthogonal tag orientations are simulated. The mesh contour is superposed on the images.	145
9	Simulated ultrasound images. The 4-chamber, 2-chamber and short-axis views are shown. The mesh contour is superposed on the images.	146
10	Template US recording (top row) and LCX simulation (bottom row) at matched time instants for an apical four chamber (a) and a short axis (b) slice. The E/M meshes (i.e. ground truth segmentation of the LCX simulation) are superimposed to both sequences as white contours. Template, simulation and, therefore, E/M geometry are aligned by construction at time zero (i.e. at ED1). From then on, the simulation runs independently of the template recording and, therefore, the E/M geometry no longer provides a valid segmentation of the latter. In the figure, this is particularly evident at regions marked by the white arrows. This is a consequence of the fact that, as due to the simulated ischemia, contraction in the E/M simulation is reduced as compared to the template.	147
11	Template MR recording (top row) and LCX simulation (bottom row) at matched time instants for a short-axis cine MR slice (a) and a short-axis tagged MR slice (b). The E/M meshes (i.e. ground truth segmentation of the LCX simulation) are superimposed to both sequences as white contours. Template, simulation and, therefore, E/M geometry are aligned by construction at time zero (i.e. at ED1). From then on, the simulation runs independently of the template recording and, therefore, the E/M geometry no longer provides a valid segmentation of the latter. In the figure, this is particularly evident at regions marked by the white arrows. This is a consequence of the fact that, as due to the simulated ischemia, contraction in the E/M simulation is reduced as compared to the template.	148
12	Illustration of the interest of the proposed method to better handle the interface between myocardium and the surrounding structures. a) template image at ES; b) simulated image using [?]; c) simulated image using our approach.	149
13	Histograms of myocardial intensities in both real and simulated MR data. Red dashed lines mark median intensities. Cine MR results are shown on the left side and tagged MR shown on the right side.	150

14	Statistical distribution for two rectangular speckle patches belonging to the blood pool and the myocardium (red and green blocks in (a)). Each ROI is rectangular with side equal to three times the size of the system PSF in the corresponding direction and comprises ~ 6800 samples. Note that, although voxel data is reported in (a) for a more intuitive visualization, histograms were computed on raw data, <i>i.e.</i> before down-sampling and scan conversion. In (b) and (c) the histograms of the RF signal. In (d) and (e) the histograms for the non-compressed envelope signal. In the fitting of the RF and envelope histograms Gaussian and Rayleigh distributions show good fit respectively for fully developed speckle regions (<i>i.e.</i> the blood pool) while generalized Gaussian distribution (GGD) [?] and K-distribution [?] were adopted for sub-Rayleigh regions (<i>i.e.</i> the tissue). Consistently with what is expected on real recordings, Gaussian and Rayleigh models correctly fit the amplitude data inside the blood pool while speckle inside the muscle follows instead more heavily tailed distributions. Note that histograms have been normalized to have unitary area in order to be comparable with the associated probability density functions.	151
15	a) Eight 3D patches used to evaluate speckle decorrelation. b) Time evolution of the correlation coefficient : in black the measurements within each block, in blue the confidence interval ($\mu \pm \sigma$). The dashed vertical line denotes block, in blue the confidence interval end of systole.	151
16	Regional strain curves at basal, mid and apical levels of the LV. We show one healthy, one dynssynchrony (LBBB) and three infarcted hearts (LCX, LAD, RCA).153	153
17	Global circumferential strain curves corresponding to the 18 simulated virtual patients.	154
18	Resampling of LV surface mesh to volumetric dense mesh.	157
19	Anatomical coordinates $r, l, c^{(i)}$ defined on the LV volumetric mesh (in (c), $c^{(i)}(\mathbf{x})$ has no definition in the surrounding grey region since it is beyond the range of the current window i).	159
20	Definition of LV local windows and an example of the window function for the red-circled window.	160
21	The resampling of the first simulation mesh into a new LV mesh defined following the radial, longitudinal and circumferential directions. The resampled mesh is shown in red color. The simulation mesh is shown in white with adjusted opacity level.	163
22	Local motion and strain quantification errors obtained by the sparse demons and the proposed algorithms (ADM). Both methods were run on the 18 virtual patients simulated in Sec. 4. For each cardiac phase, local motion and strains errors were computed for each 17 AHA segments and then averaged from all time instants, all segments and all virutal patients so to produce the given boxplots.	164
23	Local motion and strain quantification errors obtained by the the proposed ADM method using either the image intensity or the phase information as the data attachment term involved in Eq. (29). Both methods were run on the 18 synthetic tagged MR datasets simulated in Sec. 4. For each cardiac phase, local motion and strains errors were computed for each 17 AHA segments and then averaged from all time instants, all segments and all virutal patients so to produce the given boxplots.	165

1	(Figure 2 in [?]) Schematic representation of the single transmit PW method. (a) The ultrasonic array insonifies the medium using a PW transmission. (b) The backscattered RF signals are recorded by the transducer array. (c) The beam-forming procedure consists in applying time delays laws and summations to the raw RF signals to focus in the receive mode. Contrary to standard ultrasonography, each line of the image is calculated using the same RF data set but with a different set of time delays.	174
2	Principle of coherent PW compounding. The medium is insonified with steered PW, and the images are reconstructed individually for each PW. The individual images obtained from each PW are low quality images. The coherent summation of these images creates synthetic focusing throughout the complete image and allows to recover a high quality image.	175
3	(Fig. 3 in [?]) Triangle transmit sequences & motion estimation. To reduce the side effects due to sidelobes, a triangle sequence composed of 32 successive transmits is used. This figure shows an 8-transmit sequence for clarity. Slow-time autocorrelations were calculated in the ascending and descending stages, and their product was used to estimate the phase delays due to motion. The insets on the top represent the PSFs (point spread functions) with the apparent sidelobes. They rotate in clockwise then counter-clockwise directions.	176
4	Effect of the value of the steered angle α_s on the lateral resolution using synthetic acoustic field. Each pressure is compounded with 3 steered PW $(-\alpha_s, 0, \alpha_s)$ at 30-mm depth on the central axis (left column) and on the edge of the probe (right column). (a)-(b) Lateral resolution as a function of α_s . The black dashed line represents the limit of lateral resolution at 1 mm. (c)-(d) Normalized synthetic pressure computed along lateral positions for different values of α_s ($1^\circ, 5^\circ, 10^\circ, 15^\circ$). (e)-(f) 2-D normalized synthetic pressure distribution obtained for the different configurations of α_s with $5^\circ, 10^\circ$ and 15°	181
5	Influence of the number of transmitted steered PW on the lateral resolution for the conventional and the dichotomous scheme.	182
6	Normalized synthetic pressure computed along the lateral position at 30-mm depth in the center of the image with varying numbers of steered PW (3, 11 and 21) using (a) the regular approach (<i>e.g.</i> 3 steered PW : $\{-1^\circ, 0^\circ, 1^\circ\}$) and (b) the proposed dichotomous scheme (<i>e.g.</i> 3 steered PW : $\{-10^\circ, 0^\circ, 10^\circ\}$).	183
7	Schematic diagram of the CIRS tissue-mimicking phantom (model : 054GS) with the corresponding imaging planes used in the experiments	183
8	B-mode images of a CIRS phantom (model : 054GS) reconstructed from the three reconstructed methods using (a) 1-PW and (b) 21 steered PW with coherent compounding. (c) Lateral resolution measurements as a function of the number of steered PW. The solid lines correspond the proposed dichotomous transmission scheme while the dashed lines correspond to the regular scheme.	184
9	Contrast ratio CR as a function of the number of steered PW for the three reconstructed methods. The solid lines correspond the proposed dichotomous transmission scheme while the dashed lines correspond to the regular scheme. . .	185
10	Reconstruction procedure for conventional Fourier-based method. (a) the received raw-data in (x, t) space; (b) the 2D Fourier transform of the raw-data; (c) the remapped object spectrum; (d) the reconstructed image.	186
11	Illustration of a wavefront of a steered PW with direction \mathbf{n} . All the \mathbf{x} in the blue line has the same wavefront.	187

12	Illustration of the image formation introduced in UFSB theory	188
13	Reconstruction procedure with the proposed method. (a) the received raw-data in (x, t) space; (b) the 1D Fourier transform of all the projections of the raw-data at different angles ξ_i ; (c) the remapped object spectrum; (d) the reconstructed image.	190
14	Illustration of k-space sampling derived from the UFSB method using in emission either (a) a PW perpendicular to the probe or (b) a steered PW with direction φ_j	191
15	Point scatterer images reconstructed from the UFSB method using 1 PW in transmission with steered angles of (a) -10° (b) 0° and (c) 10° . (d) Compounded image of the previous 3 steered PW. The red circles correspond to the true position of the involved scatterers.	193
16	Lateral resolution as a function of the number of PW computed at 40-mm depth.	193
17	B-mode images of an anechoic cyst phantom reconstructed from the four PW imaging methods using (a) 1-PW and (b) 21-PW with coherent compounding. (c) Contrast ratio CR as a function of the number of steered PW for the 4 reconstruction methods.	195
18	Schematic diagram of the CIRS tissue-mimicking phantom with the corresponding imaging planes used in the experiments	196
19	B-mode images of a CIRS phantom reconstructed from the four reconstructed methods using (a) 1 PW and (b) 21 steered PW with coherent compounding. (c) Lateral resolution measurements as a function of the number of steered PW.	197
20	Contrast ratio CR as a function of the number of steered PW for the four PW imaging methods.	198
21	<i>In vivo</i> experiments - Carotid image reconstructed from the four PW imaging methods using (a) 1 PW and (b) 21 steered PW with coherent compounding. (c) Absolute differences of the UFSB and the other 3 reconstructed methods from the figures in (b).	199
22	Spectrum mapping obtained from the experimental images given at the top row of Fig. 19. (a) Magnitude of the spectrum reconstructed using the proposed UFSB technique. (b) Absolute differences of Lu's and Garcia's spectrum. (c) Absolute differences of Lu's and UFSB's spectrum.	200
23	Illustration of the travel time (t_e and t_r) involved when either (a) a DW or (b) PW is used in transmission. In both cases, ($t_e + t_r$) corresponds to the time needed by the insonifying wave to reach a scatterer (orange dot) and then come back to a specific transducer element.	204
24	Error map derived from the travel time equivalence between a sectorial acquisition using DW and the equivalent PW imaging after the application of the proposed spatial transformation. The computed error is expressed relatively to the wavelength used in the experiments. The black line on the top of the figure models the transducer location while the axis origin is located at the center of the probe. The two black curves drawn inside the error map represent the boundary of the region at the edge of the image where the maximum error is higher than $\lambda/8$	205
25	Schematic diagram of the CIRS tissue-mimicking phantom (model : 054GS) with the corresponding imaging sector used in the DWs experiments.	208
26	Synthetic phantom reconstructed from the extension of the Fourier-based methods of Bernard <i>et al.</i> and Lu <i>et al.</i> as well as the method of Papadacci <i>et al.</i> The red points correspond to the real position of the scatterers involved in the simulation.	209

27	Numerical phantom - lateral resolution measurements as a function of the number of DWs computed at 20-mm, 40-mm, 60-mm and 80-mm at (a) the middle of the image (azimuth angle of 0°) and (b) at the edge of the image (azimuth angle of 40°).	210
28	Approximation of the full width at half maximum of the main lobe corresponding to the coherent summation of DWs in the case of a monochromatic far-field at a focal distance of 60-mm.	211
29	Numerical phantom - (a) contrast ratio measurements computed at the middle of the image (azimuth angle of 0°) as a function of the number of DWs. B-mode images reconstructed from (b) the proposed extension of Bernard's method, (c) the proposed extension of Lu's method and (d) the spatial-based method of Papadacci using 15 DWs with coherent compounding.	212
30	Numerical phantom - (a) contrast ratio measurements computed at the edge of the image (azimuth angle of 40°) as a function of the number of DWs. B-mode images reconstructed from (b) the proposed extension of Bernard's method using 15 DWs with coherent compounding.	212
31	Experimental acquisition - lateral resolution measurements computed at the middle of the image (azimuth angle of 0°) as a function of the number of DWs computed at 20-mm, 40-mm, 60-mm and 80-mm.	213
32	Real acquisition - (a) Contrast ratio measurements as a function of the number of DWs. B-mode images reconstructed from (b) the proposed extension of Bernard's method, (c) the proposed extension of Lu's method and (d) the spatial-based method of Papadacci using 15 DWs with coherent compounding.	214
33	Motion compensation for high frame rate echocardiography of the left ventricle using the sectorial extension of Lu's method. Each frame has been reconstructed from the transmission of 32 DWs (tilted from -25° to $+25^\circ$ with a triangle strategy as described in [?]). Online movies are also available in the supplementary materials. The data from Porée <i>et al.</i> [?] were reanalyzed with the technique described in the present paper.	215

List of figures

Personal presentation

Curriculum vitae

OLIVIER BERNARD, born on October 8, 1980 at Saint-Etienne, France.

ACADEMIC POSITION :

Title : Associate professor		Engineering school : INSA-Lyon
CNU section number : 61		In place since : 1 st Sept. 2007
Laboratory : CREATIS, UMR 5220 - INSERM U1206.		Department : Electrical dept.

OFFICE ADDRESS :

✉ : INSA-Lyon
CREATIS Laboratory
bât. Léonard de Vinci
21 avenue Jean Capelle
69621 Villeurbanne Cedex (France)

☎ : +33 (0)4 72 43 75 71

@ : olivier.bernard@insa-lyon.fr

1 Training

2003 - 2006 : **PhD thesis** in signal and image processing.

CREATIS Laboratory, UMR 5220 - INSERM U1206 - INSA Lyon , France.

Topic : *Segmentation in echocardiographic imaging using parametric level set model driving by the statistics of the radiofrequency signal.*

Supervised by Pr. Denis Friboulet.

EEA (Electronics, Electrotechnics and Automatic) Doctoral School, University of Lyon, France.

2002 - 2003 : **Master of Research** in signal and image processing.

CREATIS Laboratory, UMR 5220 - INSERM U1206 - INSA Lyon, France.

Topic : *Multi-resolution and multi-parametric ultrasound image segmentation based on global and local statistics.*

Supervised by Pr. Olivier Basset. EEAP (Electronics, Electrotechnics, Automatic and Process) master of research, Systems ans Images branch, INSA-Lyon, France.

1998 - 2003 : **Engineering School** at the National Institute of Applied Science (INSA), Lyon, France (Electrical Department from 2001 to 2003).

1995 - 1998 : Baccalauréat série S, spécialité Mathématiques, mention *bien*.
Lycée François Mauriac (Andrézieux Bouthéon, 42).

2 Position - Experience

2007 - now : Associate professor

Creatis laboratory, UMR 5220 - INSERM U1206, University of Lyon, Electrical department of INSA, Lyon, France.

2014 : CNRS delegation

One year grant from the CNRS institute to perform full-time research on the topic of ultrafast ultrasound imaging.

2013 : Invited professor

Signal Processing Laboratory (LTS5), Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland.

Subject : *Fourier-based technique for ultrafast ultrasound imaging*

Invitation from Pr. Jean-Philippe Thiran.

2013 : CRCT allocation

One year grant from the French national council of universities (CNU) to perform full-time research on the topic of ultrasound image formation.

2007 : Postdoctoral fellowship

Biomedical Imaging Group, Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland.

Topic : *B-Spline formulation for level-set segmentation techniques*

Supervised by Pr. Michael Unser.

3 Editorial responsibilities

2013 - 2016 : Associate editor

International journal : IEEE Transactions on Image Processing (IEEE TIP).

2014 : Associate editor

International conferences : IEEE ICIP'14 and IEEE ISBI'14.

4 Awards

2012 - 2020 : PEDR grants

French grant (Primes d'Excellence Scientifique) to promote my research.

2014 : Special Mention (2nd prize) for best PhD in France

Awarded by the IEEE Engineering in Medicine and Biology Society.

5 Teaching activities

My teaching activities began during my thesis, during which I was temporary employee for 3 years at the undergraduate level of INSA Lyon, within the IT team. This first experience was focused on teaching basics in informatics through database management and computer programming using the Pascal language. I then got a one year associate professor position (ATER) at the Department of Electrical Engineering and Industrial Computing (IUT GEII) at the University of Lyon 1, France. This allowed me to diversify my teaching activity, particularly in digital electronics and industrial computing. Since 2007, I teach at the Electrical Engineering (GE) department of INSA Lyon, France. I teach four main themes : industrial computing (C programming language and microcontrollers), object-oriented programming (C++ programming language), signal and image processing and applied mathematics (neural networks). In parallel, I was in charge of the Systems and Images branch of the EEAP (Electronics, Electrotechnics, Automatic and Process) master of research at Lyon (France) from 2009 to 2012. During this period, I also taught advanced concepts in image processing.

This section presents in a synthetic manner my teaching activities ordered by level using the french LMD system (License, Master and Doctoral level). For each course, the topic, the number of hours per student and per year, the number of students and the corresponding period are provided. For convenience, the training sessions will be refereed as TD, while the practical sessions will be named as TP.

5.1 Summary table of my teaching activities

I give an average of **220** hours of classes per year (without taking into account the internship supervision), mostly through training and practical sessions since 2007. A summary of all my teaching activities I realized until now is provided in Table 1.

Topic	Type	Level	Hours per student
Office automation	TP	L1	10h
Initiation to database management	TP	L1	4h
Initiation to computer programming	TD/TP	L1	35h
Digital electronics	TP	L1	92h
Architecture of processor systems	TD/TP	L1	22h
Control/command of industrial appli.	TP	L1	22h
Computer programming	TP/project	L1	60h
Initiation to signal processing	Lecture/TD/TP	L2	61h
Industrial computing	TD/TP	L3	24h
Computer programming	TD/Project	L3	20h
Applied Mathematics	TD/TP	L3	37h
Signal processing	TD/TP	L3	56h
Computer programming	TD/Project	L3	20h
Industrial computing	TD/TP	M1	40h
Object-oriented programming	TD/TP	M1	24h
Image processing	Lecture/TP/Project	M2	35h
Advanced concepts in image processing	Lecture	M2	4h

TABLE 1 – Summary of my teaching activities.

5.2 Teaching activities during my PhD (147 hours in total)

– L1 level –

Office automation courses	TP	10h	2003-2006	30 students
----------------------------------	----	-----	-----------	-------------

Contents : Update student knowledge in computer science, in particular with office automation softwares (*e.g.* OpenOffice suite) and Linux operating system.

– L1 level –

Initiation to database management	TP	4h	2003-2006	30 students
--	----	----	-----------	-------------

Contents : Consolidate the students knowledge on database management (database access, query, information extraction, design of relational schemas and the corresponding database) thanks to training and practical sessions based on SQL language.

– L1 level –

Initiation to computer programming through Pascal language	TD/TP	35h	2003-2006	30 students
---	-------	-----	-----------	-------------

Contents : Consolidate the algorithmic knowledge of the student (problem solving in the form of instructions and elementary loops). The underlying concepts (functions, methods, structures) were then practiced based on Pascal language through the realization of two projects (*e.g.* classical sorting problem).

5.3 Teaching activities during my one year associate professor (ATER) position (196 hours in total)

– L1 level –

Digital Electronics and Logical Systems	TP	92h	2006-2007	30 students
--	----	-----	-----------	-------------

Contents : Update student knowledge in digital electronics : binary representation of numbers, logical operators, Karnaugh map, basics in VHDL language.

– L1 level –

Architecture of processor systems	TD/TP	22h	2006-2007	30 students
--	-------	-----	-----------	-------------

Contents : Fundamentals in assembly language programming. During the training sessions the students learned an instruction set of a specific microcontroller while the practical session allowed the implementation of simple programs.

– L2 level –

Control/command of industrial applications	TP	22h	2006-2007	30 students
---	----	-----	-----------	-------------

Contents : Basics in control/command of industrial application : normalized system representation through Grafset description language.

– L1 level –

Computer programming through C language	TP/project	60h	2006-2007	30 students
--	------------	-----	-----------	-------------

Contents : Consolidate the algorithmic knowledge of the student (problem solving in the form of instructions and elementary loops). The underlying concepts (functions, methods, structures) were then practiced based on C language (pointer, function call, input/output function parameters, array) through practical sessions and the realization of a project.

5.4 Teaching activities as a full-time associate professor at INSA-Lyon (220 hours on average per year)

– L3 level –

Industrial computing through assembly programming	TD/TP	24h	2007-2012	30 students
--	-------	-----	-----------	-------------

Contents : Fundamentals in hardware architecture of data-processing systems (overall organization, memory access, instruction set). During the training sessions, a particular microcontroller solution was studied (PIC, Microship Technology) and used in practical sessions to implement several programs.

– M1 level –

Industrial computing through assembly programming	TD/TP	40h	2007-2012	30 students
	TP	32h	2015-	15 students

Contents : Consolidate students knowledge in hardware architecture of data-processing systems (shared memory, interruption and stack) through training and practical sessions. The same microcontroller (PIC, Microship Technology) than the one introduced the year before was used during the practical sessions.

– L3 level –

Computer programming through C language	TD	16h	2007-2012	30 students
	TD/project	20h	2015-	30 students

Contents : Consolidate students knowledge in C programming language (structure, pointer, function, input/output function parameters, multi-dimensional array). Since 2015, I am deeply involved in the supervision of IT projects in relation with the applied mathematics module presented below.

– M1 level –

Object-oriented programming through C++ language	TD/TP	24h	2007-2012	30 students
	TD/TP	24h	2015-2017	30 students

Contents : Fundamentals in object-oriented programming language through UML and C++ language (class/object, attributes/methods, pointers/references, template, STL).

– L3 level –

Applied Mathematics	TD/TP	37h	2015-	30 students
----------------------------	-------	-----	-------	-------------

Contents : Consolidate student knowledge in applied mathematics : fundamental transformations used in signal theory (Laplacian, Z and Fourier), interpolation/approximation, matrix inversion through LU and Cholesky decomposition. The different concepts are studied under Matlab language during the practical sessions. Since 2018, I actively participated in the redesign of this module for the benefit of teaching basics in neural networks

– L2 level –

Initiation to signal processing	Lecture/TD/TP	28h	2015-2017	20 students
	Lecture/TD/TP	61h	2018-	20 students

Contents : Signal processing fundamentals : signal representation, system modeling, Fourier analysis and filtering. Since 2018, I am involved in the supervision of Matlab projects (three months long) in either computer vision or medical imaging.

– L3 level –

Signal processing	TD/TP	56h	2007-2012	30 students
	TP	48h	2015-	20 students

Contents : Consolidate student knowledge in signal processing : signal acquisition and modeling, fundamentals transformations (Laplacian, Z, Fourier) for system analysis, analogical and numerical filtering. Numerical filtering concepts are reinforce during practical sessions based on the programming of Digital Signal Processors (DSP).

– M2 level –

Image processing	Lecture/TP/projects	27h	2007-2012	30 students
	Lecture/TP/projects	35h	2018-	20 students

Contents : Image processing fundamentals : image representation, filtering, mathematics morphology, contour detection, segmentation, classification. During this course, I supervise several Matlab/Python projects (three months long) in either computer vision or medical imaging.

– M2 (research) level –

advanced in image processing	Lecture	4h	2007-2012	15 students
	Lecture	4h	2015-2017	15 students

Contents : Consolidate student knowledge in image segmentation through the use of active contour models.

6 Collective responsibility

6.1 Teaching

6.1.1 Responsible for the Systems and Images part of the EEAP master of research, Lyon - France

I was responsible of the Systems and Images (SI) branch of the EEAP master of research at Lyon (France) from 2009 to 2013. In particular, I was in charge of 20 students per year. My role was to guarantee the good functioning of all the courses from an administrative point of view (secretariat service), to ensure a continuous renewal of the educational program (during my term, around 30% of the courses have been renewed) and perform student recruitment each year. The responsibility of the Systems and Images part of the master took around 20% of my working time. In 2013, I got a grant to do 2 years of full research. I thus trained an assistant from 2012 to 2013 and let him the responsibility of the SI branch of the EEAP master since 2013.

6.1.2 Elected member at the board of the Electrical Engineering Department at INSA-Lyon

In 2012, I was elected to be a member for 4 years at the board of the Electrical Department at INSA-Lyon. During the first year, I participated to 9 meetings where decisions for the good functioning of the department (mainly in terms of education and department funding) have to

be taken. Because I got a grant to perform two years of full research, I had to resign from this responsibility after one year.

6.2 Research

6.2.1 Responsible of the IT team of Creatis

From 2013 to 2016 I was also responsible of the IT team of CREATIS which is composed by four 4 engineers. My main role was to reorganize the full IT service so to better integrate the engineer work into the researcher activities. This was realized by organizing the service around projects. Each researcher at CREATIS can contact at any time an engineer according to his (her) domain of expertise (software developer, GPU programming, database management, grid deployment service) for a particular need. This is then to the charge of the engineer to define the corresponding informatics needs and to structure together with the researcher a project which has to be realized as soon as possible according to the availability of the engineer. During my term, more than 20 projects were realized, such as the organization of international challenges, the deployment of internal applications through the VIP platform (online service for the deployment of applications on grid structures)¹, the development of customized software interfaces.

6.2.2 Member of the jury of one PhD defense

I was a member of jury of the PhD defense of Doctor Romane Gauriau in 2015 whose topic was : shape-based approaches for fast multi-organ localization and segmentation in 3d medical images. This PhD was done in collaboration between Telecom Paris University and Philips medisys compagny.

6.3 Associate professor recruitment jury

Since 2007, I participated to two juries for the recruitment of associate professors (named MCF for "Maître de Conférences) in France :

MCF 0077 61 section (Computer Engineering, Automation and Signal Processing), deep learning in medical imaging, Informatics teaching at INSA-Lyon, 2017.

MCF 4238 61 section (Computer Engineering, Automation and Signal Processing), deep learning in medical imaging, Informatics teaching at INSA-Lyon, 2019.

1. <https://www.creatis.insa-lyon.fr/vip/>

Supervision and scientific impacts

1 Supervision activities

Since I got my permanent position as an associate professor at INSA-Lyon, I co-supervised **8 PhD theses** (of which 5 have been already defended) and supervised **2 post-doctoral fellowships** and **15 masters of research**. The corresponding research topics are listed as follows :

1. cardiac image analysis,
2. ultrasound image simulation,
3. ultrasound image reconstruction,
4. biomedical image analysis.

A summary of all the research activities I realized since 2007 is provided in Table 1. In each case, the research topic, the year and the supervisors (names in *italic* correspond to industrial supervisors) are provided.

1.1 PhD theses

All PhDs are presented in the same way, following the framework given below :

Student name, Position after the PhD.

<i>PhD title</i>		Date of PhD defense
Type of grant	<u>Main supervisor</u> , co-supervisor	Publications IJ (international journals), IC (international conferences)

Type	Student name	Topic	Year	Co-supervisors
PhD	Thomas Dietenbeck	1	2009-2012	D. Friboulet
	Daniel Barbosa	1	2009-2013	D. Friboulet, J. D'hooge
	Miaomiao Zhang	3	2013-2016	D. Friboulet
	Yitian Zhou	1, 2	2014-2017	D. Friboulet, <i>M. De Craene</i>
	Joao Pedrosa	1	2014-2018	J. D'hooge
	Sarah Leclerc	1	2016-2019	C. Lartizien
	Yunyun Sun	1, 2	2018-2021	P. Clarysse
	Ewan Evain	1, 2	2018-2021	D. Friboulet, <i>M. De Craene</i>
Postdoctoral	Khuram Faraz	2	2018-2019	
	Martino Alessandrini	1, 2	2011-2013	H. Liebgott
Master	Thomas Dietenbeck	4	2008-2009	
	Mohamed-salah Deghiche	1	2008-2009	D. Friboulet
	Benoît Perrot	4	2009-2010	S. Valette
	Antonin Perrot-Audet	4	2009-2010	
	Kalaimaran Routtramourhy	4	2010-2011	O. Basset
	Thomas Philibert	4	2010-2011	
	Miaomiao Zhang	3	2011-2012	D. Friboulet
	Jan Hoskovec	3	2011-2012	
	Shaojie Wang	1	2014-2015	T. Grenier
	Sarah Leclerc	1	2015-2016	T. Grenier
	Mathilde Caron	1	2015-2016	<i>M. De Craene</i>
	Gaoyang Cai	3	2016-2017	B. Nicolas
	Feriel Khellaf	1	2016-2017	J. Bosh
	Yunyun Sun	1	2017-2018	
	Ewan Evain	1	2017-2018	<i>M. De Craene</i>

TABLE 1 – Summary of my research supervisory.

Thomas Dietenbeck, Associate professor, Pierre & Marie Curie University, Paris.

<i>Segmentation of 2D-echocardiographic sequences using level-set constrained with shape and motion priors.</i>		Defended on November, 29th, 2012
French government grant	<u>Denis Friboulet</u> (50%), Olivier Bernard (50%)	IJ [21, 31], IC [60, 72, 90, 93]

Contributions :

- integration of a shape prior term adapted to heart morphology into the level-set framework,
- integration of motion information into the level-set formalism to propagate the segmentation process over the whole cardiac sequence.

Summary :

Echocardiography is one of the main cardiac imaging technique, since it is non invasive and allows a real time acquisition of cardiac structures. The segmentation of the cardiac muscle (myocardium) in these sequences is thus an important goal for the analysis of the heart function and for diagnosis assistance. However this segmentation turns out to be a difficult task due to the physics of ultrasound acquisition which yield inhomogeneities inside the structures and a poor contrast between the myocardium and the surrounding issues.

In this work, the myocardium segmentation problem is tackle using the level-set formalism. The myocardium is first approximated by a geometric model (hyperquadrics) which allows to handle asymmetric shapes such as the myocardium while avoiding a learning step. This representation is then embedded into the level-set formalism as a shape prior for the joint segmentation of the endocardial and epicardial borders. This shape prior term is coupled with a local data attachment term and a thickness term that prevents both contours from merging. The algorithm is validated on a dataset of 80 images at end diastolic and end systolic phase with manual references from 3 cardiologists.

In a second step, the whole sequences is segmented using motion information. To this end, a level conservation constraint is applied on the implicit function associated to the level-set thanks to a novel energy term expressed through a variational framework. This energy is then added to the previously described algorithm in order to constrain the temporal evolution of the contour. Finally the algorithm is validated on 20 echocardiographic sequences with manual references of 2 experts (corresponding to approximately 1200 images).

Daniel Barbosa, postdoctorat fellow, Portugal.

<i>Automated assessment of cardiac morphology and function : an integrated B-Spline framework for real-time segmentation and tracking of the left ventricle.</i>		Defended on October, 28th, 2013
European grant	Jan d'hooge (33%), Denis Friboulet (33%), Olivier Bernard (33%)	IJ [6, 20, 26, 29, 30], IC [66, 69, 70, 75, 81, 82, 86, 88, 89, 96]

Contributions :

- development of a new segmentation paradigm named B-Spline Explicit Active Surface (BEAS) specifically designed to segment 3D shapes near real time,
- strong validation of the application of the BEAS formalism for real-time segmentation of the left ventricle in 3D echocardiography.

Summary :

With the increasing prevalence of cardiovascular diseases, the need for advanced diagnosis systems that are able to detect early cardiac dysfunction are now needed more than ever. Real-time 3D echocardiography has made its way into clinical practice over the last decade and is now generally accepted as a competitive alternative to cardiac magnetic resonance imaging for volumetric assessment of left ventricular morphology and function. Nonetheless, there is still a need for software tools enabling a faster, more accurate analysis, while reducing the burden to the operating physician and minimizing the intra and interobserver variability of the measured indices.

The fundamental methodological innovation presented directly focuses on the inheritance of desirable properties of level-set oriented algorithms, such as advanced region-based segmentation energies and fast/robust interface evolution via B-Spline filtering, while dramatically reducing the computational load associated with 3D segmentation problems. This was possible through a B-spline formulation of the original Active Geometric Functions framework, and by further exploring the mathematical link between explicit and implicit formulations for the image segmentation problem.

The resulting algorithm provides a very competitive balance between accuracy and computational burden, enabling real-time 3D segmentation applications. Building on this technical breakthrough, we extensively validated its use for left ventricle volumetric assessment in a clinical setting, while at the same time dealing with some fundamental limitations such as its initialization, the user interaction with the segmented surface and the integration of temporal information in the boundary identification and tracking problems. This results in an coordinated suite of algorithms targeting real-time, fully automatic segmentation and tracking of the left ventricle during the cardiac cycle.

Miaomiao Zhang, Postdoctoral fellow, China.

<i>Fourier-based reconstruction of ultrafast sectorial images in ultrasound.</i>		Defended on December, 16th, 2016
Chinese government grant	<u>Denis Friboulet</u> (50%), Olivier Bernard (50%)	IJ [10], IC [47, 48, 51, 58, 64]

Contributions :

- development of a novel Fourier-based plane wave imaging method based on the Fourier slice theorem used in tomography,
- extension of 2D Fourier-based imaging methods from linear to sectorial imaging,
- extension of proposed Fourier-based diverging wave imaging method to 3D.

Summary :

Echocardiography is one of the most widely used modality in real time heart imaging thanks to its noninvasive nature and low cost. By providing dense image volumes in real time, three-dimensional echocardiography can improve the accuracy of the echocardiographic evaluation of cardiac chamber volumes and be used to assess cardiovascular function and anatomy in various clinical settings. However, the real-time property of three-dimensional echocardiography is still limited in terms of frame rate due to the limited speed of sound. To increase the frame rate, plane wave and diverging wave in transmission have been proposed to drastically reduce the number of transmissions to reconstruct one image. In this thesis, starting with the 2D plane wave imaging methods, the reconstruction of 2D/3D echocardiographic sequences in Fourier domain using diverging waves is addressed. The main contributions are described below.

The first contribution concerns the development of an alternative Fourier-based plane wave imaging method (*i.e.* Ultrasound Fourier Slice Beamforming) by using the concept of steered plane waves both in transmission and in reception. We build a theoretical model to describe the relationship between the echoes and object function based on the Fourier slice theorem. Results revealed that the proposed method produces very competitive image quality compared to the state-of-the-art Fourier-based and spatial-based methods.

The second contribution concerns the extension of Fourier-based imaging methods from linear to sectorial imaging in 2D by studying the difference between plane wave and diverging wave transmission in terms of travel time for a given scatterer in the medium and a given transducer element. We derive an explicit spatial transformation which allows deforming the referential Cartesian space insonified by a diverging wave into a dedicated one where the modified medium can be considered as being excited by a plane wave. Comparisons with the current state-of-the-art method illustrate the potential of the derived methods in producing competitive results with lower computational complexity when compared to the conventional delay and sum (DAS) technique.

Finally, the 2D Fourier-based diverging wave imaging methods are extended to 3D by deriving the equivalence between 3D diverging wave and plane wave. Results show that the proposed approach provides competitive scores in terms of image quality compared to the DAS technique, but with a much lower computational complexity.

Yitian Zhou, Postdoctorat fellow, France.

<i>Quantification du mouvement et de la déformation cardiaques à partir d'IRM marquée tridimensionnelle sur des données acquises par des imageurs Philips.</i>		Defended on July, 03rd, 2017
Industrial grant (CIFRE Philips)	<u>Denis Friboulet</u> (10%), Olivier Bernard (45%), Mathieu De Craene (45%, Philips)	IJ [3, 17], IC [49, 53, 54]

Contributions :

- development of a fast 3D tagged MR quantification algorithm,
- design of a novel pipeline for generating 3D synthetic cardiac US and MR image sequences for validation purposes.

Summary :

Cardiovascular disease is one of the major causes of death worldwide. A number of heart diseases, such as hypertrophy, dilated cardiomyopathy and myocardial infarction, can be diagnosed through the analysis of cardiac images after quantifying shape and function. Recently, there is a surge in the development of fast 3D cardiac imaging techniques in both ultrasound (US) and magnetic resonance (MR) imaging, making it possible to quantify myocardial motion and strain fully in 3D. However, the application of these deformation quantification algorithms in clinical routine is somewhat held back by the lack of a solid validation. These quantification algorithms need to be thoroughly validated before being used in clinics. The main contributions are described below.

First, a novel 3D extension of the well-known harmonic phase tracking method was derived. The point-wise phase-based optical flow tracking was combined with an anatomical regularization model in order to estimate anatomically coherent myocardial motions. In particular, special efforts were made to ensure a reasonable radial strain estimation by enforcing myocardial incompressibility through the divergence theorem. The proposed HarPAR algorithm was evaluated on both healthy volunteers and patients having different levels of ischemia. On volunteer data, the tracking accuracy was found to be as accurate as the best candidates of a recent benchmark. On patient data, strain dispersion was shown to correlate with the extent of transmural fibrosis. Besides, the ischemic segments were distinguished from healthy ones from the strain curves.

Second, a simulation pipeline for generating realistic synthetic cardiac US, cine and tagged MR sequences from the same virtual subject was designed. Template sequences, a state-of-the-art electro-mechanical (E/M) model and physical simulators were combined in a unified framework. The E/M model was exploited for simulating groundtruth cardiac motion fields. The template sequences were registered to the simulations by a novel warping technique aimed at ensuring a synthetic motion consistent with the E/M model and a smooth transition between the myocardium and the background. Finally, backscattering amplitudes and effective proton densities were derived from the warped templates respectively for US and MR simulations to exploit the corresponding physical simulators for generating image data. In total, we simulated 18 virtual patients (3 healthy, 3 dyssynchrony and 12 ischemia), each with synthetic sequences of 3D cine MR, US and tagged MR. The synthetic images were assessed both qualitatively and quantitatively. They showed realistic image textures similar to real acquisitions. Besides, both the ejection fraction and regional strain values are in agreement with reference values published in the literature.

Joao Pedrosa, Postdoctorat fellow, Portugal.

<i>Left ventricular segmentation of the heart in real-time 3D echocardiography.</i>		Defended on October, 10th, 2018
European grant	Jan D’hooge (80%), Olivier Bernard (20%)	IJ [5, 6, 14], IC [45]

Contributions :

- extension of an existing B-spline Explicit Active Surfaces framework to efficiently combine segmentation and tracking strategies for accurate full cycle left ventricle segmentation,
- efficient integration of statistical shape model into the B-spline Explicit Active Surfaces framework to improve the robustness of the left ventricle segmentation.

Summary :

Echocardiography plays a crucial role in clinical cardiology with diagnostic, prognostic and interventional value. Though 2D echocardiography remains the predominant modality for cardiac assessment, 3D echocardiography has the advantage of allowing for a 3D rendering of the anatomical structures thus discarding the need for the geometrical assumptions inherent to 2D echocardiography. Nevertheless, the analysis of 3D images is challenging and time consuming, fueling the need for software tools that enable a fast, accurate analysis that reduces the burden on the clinician while at the same time reducing the inter- and intra observer variability of the clinical indices extracted. In this context, the main contributions of this PhD are described below.

First, based on the B-spline Explicit Active Surfaces framework, a hybrid energy for segmentation and tracking was introduced, allowing for more accurate full cycle segmentation. Robust left ventricular myocardial segmentation was reinforced by an efficient coupling between the endo- and epicardial surfaces. Moreover, automatic short axis orientation of the left ventricle was proposed, allowing for additional information to be extracted.

Second, a statistical shape model built from cardiac magnetic resonance imaging was used to improve the robustness of the segmentation by providing information on the expected shapes of the left ventricle when image information is low or unreliable.

At each step of the development thorough validation of the methods was performed, leading to the final framework for fast, automatic and robust full heart cycle 3D left ventricular myocardial segmentation. This framework was then implemented in a user-friendly distributable software application to potentiate its application in future studies. The added value of 3D echocardiography was further highlighted through a method for myocardial performance mapping through the study of the stress-strain relationships validated against nuclear imaging.

Sarah Leclerc, PhD student, France.

<i>Multi-structural segmentation in cardiac imaging by the use of machine learning.</i>		Defense scheduled for October 2019.
French government grant	<u>Carole Lartizien</u> (50%), Olivier Bernard (50%)	IJ [1], IC [41-44]

Contributions :

- development of the largest publicly-available and fully-annotated dataset for 2D echocardiographic assessment (to our knowledge),
- evaluate how far the state-of-the-art U-Net deep convolutional network method can go at assessing 2D echocardiographic images, *i.e.* segmenting cardiac structures as well as estimating clinical indices, thanks to the database especially designed to answer this objective.

Summary :

Analysis of 2D echocardiographic images plays a crucial role in clinical routine to measure the cardiac morphology and function and to reach a diagnosis. Such analysis is based on the interpretation of clinical indices which are extracted from low-level image processing such as segmentation and tracking. In clinical routine, semi-automatic or manual annotation is still daily work due to the lack of accuracy and reproducibility of fully-automatic cardiac segmentation methods. This leads to time consuming tasks prone to intra- and inter-observer variability. In this context, the main contributions of this PhD are described below.

First, a new public dataset called CAMUS (Cardiac Acquisitions for Multi-structure Ultrasound Segmentation) was setup. This dataset contains 2D echocardiographic sequences (both two and four-chamber views) of 500 patients that were acquired in the same medical center with the same equipment. The size of this dataset as well as its tight bond to every-day clinical issues give the possibility to train deep learning methods to automatically analyze echocardiographic data.

Thanks to this dataset, a fine-tuned version of the well-known U-Net convolutional neural network (CNN) was implemented and evaluated. Results demonstrated that this CNN method can provide highly accurate segmentation results compared to state-of-the-art techniques within the inter-observer variability, allowing a robust and reliable automatic estimation of the end diastolic and end systolic volumes of the left ventricle. However, the developed U-Net method still fails at accurately estimating the the ejection fraction of the left ventricle, which illustrates the need to integrate more advanced concepts into conventional U-Net architectures.

The last year of the PhD is thus focused on the introduction of such concepts into the U-Net framework so to better extract the ejection fraction of the left ventricle and to decrease the number of anatomically incoherent segmentation at the same time.

Yunyun Sun, PhD student, France.

<i>Robust estimation of hemodynamic clinical indices in ultrafast ultrasound imaging based on deep learning</i>		Defense scheduled for October 2021.
French government grant	<u>Patrick Clarysse</u> (50%), Olivier Bernard (50%)	

Summary :

Recent advances in ultrafast ultrasound imaging allows the derivation of new clinical indices to diagnose cardiac pathologies. In this context, the estimation of the vortex dynamic represents a novel index with strong potential to characterize the cardiac function. However, ultrafast ultrasound imaging suffers from an inherent lower signal to noise ratio compared to conventional imaging, which inevitably impacts the quality of the extracted measurements.

In this context, the objective of this thesis is to make the estimation of such indices more robust thanks to the development of dedicated deep learning architectures. The targeted contributions of this Phd are the following :

1. numerical modeling of the intraventricular vectorial flow based on an Lagrangian formulation where the fluid will be represented as a set of particles interacting with each others and the myocardial interface,
2. integration of the numerical modeling output into a physical ultrasound simulation pipeline to generate realistic synthetic sequences,
3. development of a supervised deep learning architecture to estimate vortex dynamics based on a cohort of virtual patients generated thanks to the pipeline described above.

Ewan Evain, PhD student, France.

<i>Synthetic imaging and deep learning for robust dense field estimation in echocardiography : application to myocardial strain quantification and blood flow estimation.</i>	Defense scheduled for October 2021.
Industrial grant (CIFRE Philips)	<u>Denis Friboulet</u> (20%), Olivier Bernard (40%), Mathieu De Craene (40%, Philips)

Summary : While echocardiography is one of the mostly used imaging techniques in clinical routines, this modality still suffers from a lack of reproducibility whose reasons are complex : patient position, breathing, probe orientation and pressure, choice of the acquisition view. All these factors make difficult the acquisition of coherent images required to assess comparable measurements between experts. The establishment of databases with references remains a key challenge in ultrasound imaging both for validation and learning purposes.

In this context, the main objective of this thesis is to develop deep learning architectures dedicated to the quantification of motion in echocardiographic imaging. The targeted contributions of this PhD are the following :

1. development of a supervised deep learning architecture dedicated to the estimation of cardiac deformation in ultrasound imaging,
2. development of a cohort of virtual patients which will serve as reference to the deep learning method described above,
3. pre-clinical study to assess the contribution of the developed technique to the quantification of myocardial deformation to improve the detection of patients suffering from heart failure and showing abnormalities of filling (with or without abnormal systolic function).

1.2 Masters of research

- 2008-2009** : Thomas Dietenbeck (Master INSA-Lyon), *Segmentation through local B-Spline level-set model.*
- 2008-2009** : Mohamed Salah Deghiche (Master INSA-Lyon), *Segmentation of echocardiographic images through active contours with shape constraints.*
- 2009-2010** : Antonin Perrot-Audet (Master INSA-Lyon), *Localization and segmentation of cells in 3D confocal microscopy imaging.*
- 2009-2010** : Benoît Perrot (Master Strasbourg University), *Segmentation with user constraints.*
- 2010-2011** : Thomas Philibert (Master INSA-Lyon), *Automatic detection and segmentation of brain aneurysm in CT imaging.*
- 2010-2011** : Kalaimaran Routtramourhy (Master INSA-Lyon), *Characterization of biologic tissue through ultrasound signal deconvolution.*
- 2011-2012** : Jan Hoskovec (Master INSA-Lyon), *Ultrasound image reconstruction through Fourier space.*
- 2011-2012** : Miaomiao Zhang (Master INSA-Lyon), *Ultrasound image reconstruction using distributed compressive sensing.*
- 2014-2015** : Shaojie Wang (Master INSA-Lyon), *Cardiac structure detection on ultrasound 3-D image using structured random forest.*
- 2015-2016** : Sarah Leclerc (Master INSA-Lyon), *Echocardiography image segmentation through structured random forest.*
- 2015-2016** : Mathilde Caron (Master École polytechnique-Paris), *Cardiac motion and strain estimation with anatomical regularization.*
- 2016-2017** : Gaoyang Cai (Master INSA-Lyon), *Speed estimation of ultrasound wave in ultrafast imaging.*
- 2016-2017** : Ferial Khellaf (Master Polytech-Lyon), *Left ventricle segmentation in 3D ultrasound by structured random forests and active shape models.*
- 2017-2018** : Yunyun Sun (Master Telecom Saint-Etienne), *Echocardiography image segmentation through deep learning.*
- 2017-2018** : Ewan Evain (Master ECE / UPMC Paris), *Robust estimation of myocardial muscle deformation in echocardiographic imaging by deep neural network.*

2 Scientific impacts

2.1 International collaborations

Since the beginning of my tenure as an associate professor, I set up several international collaborations to reinforce my research projects :

- collaboration with Pr. **Jan D’hooge** (KU Leuven, Belgium) on the segmentation of ultrasound images through the B-Spline Explicit Active Surface formalism. We co-supervised 3 PhD theses (Thomas Dietenbeck, Daniel Barbosa and Joao Pedrosa) which were valued by 9 publications **IJ**[5, 6, 14, 20, 21, 26, 29–31] and 15 conferences **IC**[45, 60, 66, 69, 70, 72, 75, 81, 82, 86, 88–90, 93, 96],
- collaboration with Pr. **Jean-Philippe Thiran** (EPFL, Switzerland) on the reconstruction of ultrafast ultrasound images. This works led to the publication of 2 journal articles **IJ**[9, 10] and 7 conferences **IC**[47, 48, 51, 52, 55, 58, 64]
- collaboration with Pr. **Lasse Lovstakken** (NTNU, Norway) on the segmentation of echocardiography through deep learning. This works led to the publication of 1 journal article **IJ**[1] and 2 conferences **IC**[41, 42],
- collaboration with Pr. **Pierre-Marc Jodoin** (Sherbrook University, Canada) on the segmentation of cardiac imaging through deep learning. This works led to the publication of 2 journal articles **IJ**[1, 4] and 1 conference **IC**[41].

2.2 National collaboration

I work with Dr. **Mathieu De Craene** from Philips Medisys (France) on the simulation of realistic cardiac sequences for validation purposes since 2013. In this regard, we co-supervised the PhD thesis of Yitian Zhou between 2014 and 2017. This collaboration was valued by 6 publications **IJ**[2, 3, 12, 17, 18, 24] and 5 conferences **IC**[49, 53, 54, 59, 65].

2.3 Awards and distinctions

- Special Mention (2nd prize) for best PhD in France awarded by the IEEE Engineering in Medicine and Biology Society.
- One year grant from the CNRS institute to perform full-time research on the topic of ultrafast ultrasound imaging.
- Invited professor at the signal processing laboratory (LTS5), Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland. During this stay, I worked with Pr. Jean-Philippe Thiran on the reconstruction of ultrafast ultrasound images through the Fourier space.
- Five invited talks in international conferences (one of which I am first author).

2.4 Associate editor positions

I got different associate editor positions since 2012 :

- 2012-2016 : associate editor for the *IEEE Transactions on Image processing Journal* (impact factor 5.071). I was in charge of more than one hundred papers.
- 2014 : associate editor for the *IEEE International Symposium on Biomedical Imaging (ISBI’14)*. I was in charge of around 20 papers.
- 2014 : associate editor for the *IEEE International Conference on Image Processing (ICIP’2014)*. I was in charge of around 20 papers.

2.5 Reviewing work

I am a regular reviewer for the following international journals :

- *IEEE Transactions on Image Processing* (impact factor 5.071),
- *IEEE Transactions on Medical Imaging* (impact factor 6.131),
- *Medical Image Analysis* (impact factor 5.356),
- *IEEE International Ultrasonics Symposium (IUS)*,
- *IEEE International Symposium on Biomedical Imaging (ISBI)*,
- *IEEE International Conference on Image Processing (ICIP)*.

2.6 Organization of international challenges

I co-organized several international challenges since 2014 :

- Challenge on Endocardial Three-dimensional Ultrasound Segmentation (CETUS challenge) co-organized with prof. D'hooge (KU Leuven, Belgium) and Prof. Bosch (Erasmus University, the Netherlands) during MICCAI'14 conference. 9 teams participated. An open access database consisting in 45 3D ultrasound sequences with manual annotations (3D meshes of the left ventricle) has been set up for this event. The database and the online evaluation platform are still available and will be maintained and kept open as long as the data remains relevant for clinical research (www.creatis.insa-lyon.fr/EvaluationPlatform/CETUS/). This event has been valued by the publication of 2 journal articles **IJ**[7, 15] and 2 conference papers **IC**[67, 68].
- Plane-wave Imaging Challenge in Medical UltraSound (PICMUS challenge) co-organized with Prof. Liebgott (University of Lyon, France), Dr. Morales (NTNU, Norway) and Prof. Jensen (Technical University of Denmark) during the IEEE IUS'16 conference. 29 teams participated. An open access database consisting in numerical phantoms, in vitro and in vivo data has been set up for this event (<https://www.creatis.insa-lyon.fr/EvaluationPlatform/picmus/>). This event has been valued by the publication 1 conference paper **IC**[50].
- Automated Cardiac Diagnosis Challenge (ACDC challenge) co-organized with Prof. Jodoin (University of Sherbrook, Canada) and Dr. Lalande (University of Burgundy, France) during MICCAI'17 conference. 10 teams participated. An open access database consisting in 150 multi-slice 2D cine MRI sequences with manual annotations (3D binary volumes of the left ventricle, myocardium and right ventricle) has been set up for this event. The database and the online evaluation platform are still available and will be maintained and kept open as long as the data remains relevant for clinical research (<https://acdc.creatis.insa-lyon.fr/>). This event has been valued by the publication of 1 journal article **IJ**[4].

2.7 Regular involvement to scientific events

I was involved either as a lecturer or as co-organizer of several winter/spring schools :

April 2013 : talk on "Introduction in ultrasound image segmentation" during the Marie-Curie Initial Training Network USART spring-school event.

February 2017 : talk on "Modeling of ultrasound waves and image reconstruction" during the Marie-Curie Actions European Industrial Doctorate Cardiofunxion winter-school event - <https://www.youtube.com/watch?v=rKfnAZo1tKo&t=4363s>.

February 2018 : co-organizer and lecturer on "Multimodal generation of realistic synthetic images" during the Marie-Curie Actions European Industrial Doctorate Cardiofunction winter-school event - <https://www.youtube.com/watch?v=r09yCA7Aj28&t=316s>
<https://www.youtube.com/watch?v=07jiLMAsgW4>.

April 2019 : co-organizer of the french LABEX PRIMES spring-school event on deep learning for medical imaging - <https://deepimaging2019.sciencesconf.org/>.

April 2019 : co-organizer and lecturer on "Simulation of realistic ultrasound images" during the IEEE UFFC spring-school event on ultrasound motion imaging and experimentations - <http://www.biomecardio.com/UFFCschool2019/index.html>.

I also co-organized a one day seminar on neural network sponsored by the french club EEA (Electronics, Electrotechnics and Automatic) - <https://rnn2018lyon.sciencesconf.org>. During this seminar, more than 70 persons assisted to several lectures in the morning on the basics of neural networks and participated to an hands-on session in the afternoon to study a practical implementation of neural network through a given Matlab code.

2.8 Publications and communications

I have currently published 40 articles in international journals, 72 articles in international conferences and 2 book chapters. I also participated in 5 invited talks. Among my articles in international journals, 90% belong to the 1st quartile of the most cited journals in their field (Source Scimago Journal & Country Rank). In particular, I published 12 papers in IEEE Transactions on Medical Imaging, 8 papers in IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, 5 papers in Medical Image Analysis and 4 papers in IEEE Transactions on Image Processing journals.

Peer-reviewed journals

I published **40** articles in international journals.

- [1] S. Leclerc, E. Smistad, J. Pedrosa, A. Østvik, F. Cervenansky, F. Espinosa, T. Espeland, E. A. R. Berg, P. Jodoin, T. Grenier, C. Lartizien, J. D'hooge, L. Lovstakken, and **O. Bernard**. Deep learning for segmentation using an open large-scale dataset in 2d echocardiography. *IEEE Transactions on Medical Imaging*, pages 1–1, 2019.
- [2] M. Alessandrini, B. Chakraborty, B. Heyde, **O. Bernard**, M. De Craene, M. Sermesant, and J. D'Hooge. Realistic vendor-specific synthetic ultrasound data for quality assurance of 2-d speckle tracking echocardiography : Simulation pipeline and open access database. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 65(3) :411–422, March 2018.
- [3] Y. Zhou, S. Giffard-Roisin, M. De Craene, S. Camarasu-Pop, J. D'Hooge, M. Alessandrini, D. Friboulet, M. Sermesant, and **O. Bernard**. A Framework for the Generation of Realistic Synthetic Cardiac Ultrasound and Magnetic Resonance Imaging Sequences From the Same Virtual Patients. *IEEE Transactions on Medical Imaging*, 37(3) :741–754, March 2018.
- [4] **O. Bernard**, A. Lalande, C. Zotti, F. Cervenansky, X. Yang, P. Heng, I. Cetin, K. Lekadir, O. Camara, M. A. G. Ballester, G. Sanroma, S. Napel, S. Petersen, G. Tziritas, E. Grinias, M. Khened, V. A. Kollerathu, G. Krishnamurthi, M. Rohé, X. Pennec, M. Sermesant,

- F. Isensee, P. Jäger, K. H. Maier-Hein, C. F. Baumgartner, L. M. Koch, J. M. Wolterink, I. Išgum, Y. Jang, Y. Hong, J. Patravali, S. Jain, O. Humbert, and P. M. Jodoin. Deep Learning Techniques for Automatic MRI Cardiac Multi-structures Segmentation and Diagnosis : Is the Problem Solved? *IEEE Transactions on Medical Imaging*, 37(11) :2514–2525, 2018.
- [5] J. Pedrosa, S. Queirós, **O. Bernard**, J. Engvall, T. Edvardsen, E. Nagel, and J. D’hooge. Fast and Fully Automatic Left Ventricular Segmentation and Tracking in Echocardiography Using Shape-Based B-Spline Explicit Active Surfaces. *IEEE Transactions on Medical Imaging*, 36(11) :2287–2296, Nov 2017.
- [6] D. Barbosa, J. Pedrosa, B. Heyde, T. Dietenbeck, D. Friboulet, **O. Bernard**, and J. D’hooge. heartBEATS : A hybrid energy approach for real-time B-spline explicit active tracking of surfaces. *Computerized Medical Imaging and Graphics*, 62 :26 – 33, 2017.
- [7] A. Papachristidis, E. Galli, M. L. Geleijnse, B. Heyde, M. Alessandrini, D. Barbosa, M. Papitsas, G. Pagnano, K. C. Theodoropoulos, S. Zidros, E. Donal, M. J. Monaghan, **O. Bernard**, J. D’hooge, and J. G. Bosch. Standardized Delineation of Endocardial Boundaries in Three-Dimensional Left Ventricular Echocardiograms. *Journal of the American Society of Echocardiography*, 30(11) :1059 – 1069, 2017.
- [8] E. Turquin, L. Petrusca, **O. Bernard**, M. Viallon, H. Liebgott, and F. Varray. Local Orientation Imaging for Tissue Structure Using Ultrasound Imaging. *IRBM*, 38(5) :298 – 303, 2017.
- [9] A. Besson, M. Zhang, F. Varray, H. Liebgott, D. Friboulet, Y. Wiaux, J. Thiran, R. E. Carrillo, and **O. Bernard**. A Sparse Reconstruction Framework for Fourier-Based Plane-Wave Imaging. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 63(12) :2092–2106, Dec 2016.
- [10] M. Zhang, F. Varray, A. Besson, R. E. Carrillo, M. Viallon, D. Garcia, J. Thiran, D. Friboulet, H. Liebgott, and **O. Bernard**. Extension of Fourier-Based Techniques for Ultrafast Imaging in Ultrasound With Diverging Waves. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 63(12) :2125–2137, Dec 2016.
- [11] F. Varray, **O. Bernard**, S. Assou, C. Cachard, and D. Vray. Hybrid Strategy to Simulate 3-D Nonlinear Radio-Frequency Ultrasound Using a Variant Spatial PSF. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 63(9) :1390–1398, Sept 2016.
- [12] M. Alessandrini, B. Heyde, S. Queirós, S. Cygan, M. Zontak, O. Somphone, **O. Bernard**, M. Sermesant, H. Delingette, D. Barbosa, M. De Craene, M. O’Donnell, and J. D’hooge. Detailed Evaluation of Five 3D Speckle Tracking Algorithms Using Synthetic Echocardiographic Recordings. *IEEE Transactions on Medical Imaging*, 35(8) :1915–1926, Aug 2016.
- [13] N. Almeida, D. Friboulet, S. I. Sarvari, **O. Bernard**, D. Barbosa, E. Samset, and J. D’hooge. Left-Atrial Segmentation From 3-D Ultrasound Using B-Spline Explicit Active Surfaces With Scale Uncoupling. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 63(2) :212–221, Feb 2016.

- [14] J. Pedrosa, D. Barbosa, N. Almeida, **O. Bernard**, J. Bosch, and J. D’hooge. Cardiac chamber volumetric assessment using 3D ultrasound - a review. *Current Pharmaceutical Design*, 22(1) :105–121, 2016.
- [15] **O. Bernard**, J. G. Bosch, B. Heyde, M. Alessandrini, D. Barbosa, S. Camarasu-Pop, F. Cervenansky, S. Valette, O. Mirea, et al. Standardized Evaluation System for Left Ventricular Segmentation Algorithms in 3D Echocardiography. *IEEE Transactions on Medical Imaging*, 35(4) :967–977, 2016.
- [16] O. Lorintiu, H. Liebgott, M. Alessandrini, **O. Bernard**, and D. Friboulet. Compressed Sensing Reconstruction of 3D Ultrasound Data Using Dictionary Learning and Line-Wise Subsampling. *IEEE Transactions on Medical Imaging*, 34(12) :2467–2477, Dec 2015.
- [17] Y. Zhou, **O. Bernard**, E. Saloux, A. Manrique, P. Allain, S. Makram-Ebeid, and M. De Craene. 3D harmonic phase tracking with anatomical regularization. *Medical Image Analysis*, 26(1) :70–81, Dec 2015.
- [18] M. Alessandrini, M. De Craene, **O. Bernard**, S. Giffard-Roisin, P. Allain, I. Waechter-Stehle, J. Weese, E. Saloux, H. Delingette, M. Sermesant, and J. D’hooge. A Pipeline for the Generation of Realistic 3D Synthetic Echocardiographic Sequences : Methodology and Open-Access Database. *IEEE Transactions on Medical Imaging*, 34(7) :1436–1451, July 2015.
- [19] M. Alessandrini, A. Basarab, L. Boussel, X. Guo, A. Serusclat, D. Friboulet, D. Kouamé, **O. Bernard**, and H. Liebgott. A New Technique for the Estimation of Cardiac Motion in Echocardiography Based on Transverse Oscillations : A Preliminary Evaluation In Silico and a Feasibility Demonstration In Vivo. *IEEE Transactions on Medical Imaging*, 33(5) :1148–1162, May 2014.
- [20] D. Barbosa, B. Heyde, M. Cikes, T. Dietenbeck, P. Claus, D. Friboulet, **O. Bernard**, and J. D’hooge. Real-time 3D interactive segmentation of echocardiographic data through user-based deformation of B-spline explicit active surfaces. *Computerized Medical Imaging and Graphics*, 38(1) :57 – 67, 2014.
- [21] T. Dietenbeck, D. Barbosa, M. Alessandrini, R. Jasaityte, V. Robesyn, J. D’hooge, D. Friboulet, and **O. Bernard**. Whole myocardium tracking in 2D-echocardiography in multiple orientations using a motion constrained level-set. *Medical Image Analysis*, 18(3) :500 – 514, 2014.
- [22] S. Queirós, D. Barbosa, B. Heyde, P. Morais, J. L. Vilaça, D. Friboulet, **O. Bernard**, and J. D’hooge. Fast automatic myocardial segmentation in 4D cine CMR datasets. *Medical Image Analysis*, 18(7) :1115 – 1131, 2014.
- [23] J. Richy, D. Friboulet, A. Bernard, **O. Bernard**, and H. Liebgott. Blood Velocity Estimation Using Compressive Sensing. *IEEE Transactions on Medical Imaging*, 32(11) :1979–1988, Nov 2013.
- [24] M. De Craene, S. Marchesseau, B. Heyde, H. Gao, M. Alessandrini, **O. Bernard**, G. Piella, A. R. Porras, L. Tautz, A. Hennemuth, A. Prakosa, H. Liebgott, O. Somphone, P. Allain, S. Makram Ebeid, H. Delingette, M. Sermesant, J. D’hooge, and E. Saloux. 3D Strain Assessment in Ultrasound (Straus) : A Synthetic Comparison of Five Tracking Methodologies. *IEEE Transactions on Medical Imaging*, 32(9) :1632–1646, Sept 2013.

-
- [25] M. Alessandrini, A. Basarab, H. Liebgott, and **O. Bernard**. Myocardial Motion Estimation From Medical Images Using the Monogenic Signal. *IEEE Transactions on Image Processing*, 22(3) :1084–1095, March 2013.
- [26] D. Barbosa, B. Heyde, T. Dietenbeck, H. Houle, D. Friboulet, **Bernard, O.**, and J. D’hooge. Quantification of left ventricular volume and global function using a fast automated segmentation tool : validation in a clinical setting. *The International Journal of Cardiovascular Imaging*, 29(2) :309–316, Feb 2013.
- [27] T. Glatard, C. Lartizien, B. Gibaud, R. Ferreira da Silva, G. Forestier, F. Cervenansky, M. Alessandrini, H. Benoit-Cattin, **O. Bernard**, S. Camarasu-Pop, N. Cerezo, P. Clarysse, A. Gaignard, P. Hugonnard, H. Liebgott, S. Marache, A. Marion, J. Montagnat, J. Tabary, and D. Friboulet. A Virtual Imaging Platform for Multi-Modality Medical Image Simulation. *IEEE Transactions on Medical Imaging*, 32(1) :110–118, Jan 2013.
- [28] M. Alessandrini, **O. Bernard**, A. Basarab, and H. Liebgott. Multiscale optical flow computation from the monogenic signal. *IRBM*, 34(1) :33 – 37, 2013. Digital Technologies for Healthcare.
- [29] D. Barbosa, T. Dietenbeck, B. Heyde, H. Houle, D. Friboulet, J. D’hooge, and **O. Bernard**. Fast and Fully Automatic 3-D Echocardiographic Segmentation Using B-Spline Explicit Active Surfaces : Feasibility Study and Validation in a Clinical Setting. *Ultrasound in Medicine and Biology*, 39(1) :89 – 101, 2013.
- [30] D. Barbosa, T. Dietenbeck, J. Schaerer, J. D’hooge, D. Friboulet, and **O. Bernard**. B-Spline Explicit Active Surfaces : An Efficient Framework for Real-Time 3-D Region-Based Segmentation. *IEEE Transactions on Image Processing*, 21(1) :241–251, Jan 2012.
- [31] T. Dietenbeck, M. Alessandrini, D. Barbosa, J. D’hooge, D. Friboulet, and **O. Bernard**. Detection of the whole myocardium in 2D-echocardiography for multiple orientations using a geometrically constrained level-set. *Medical Image Analysis*, 16(2) :386 – 401, 2012.
- [32] J. Vandemeulebroucke, **Bernard, O.**, S. Rit, J. Kybic, P. Clarysse, and D. Sarrut. Automated segmentation of a motion mask to preserve sliding motion in deformable registration of thoracic CT. *Medical Physics*, 39(2) :1006–1015, 2012.
- [33] M. Alessandrini, S. Maggio, J. Poree, L. De Marchi, N. Speciale, E. Franceschini, **O. Bernard**, and O. Basset. A restoration framework for ultrasonic tissue characterization. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 58(11) :2344–2360, November 2011.
- [34] M. Alessandrini, T. Dietenbeck, O. Basset, D. Friboulet, and **O. Bernard**. Using a geometric formulation of annular-like shape priors for constraining variational level-sets. *Pattern Recognition Letters*, 32(9) :1240 – 1249, 2011.
- [35] B. Touil, A. Basarab, P. Delachartre, **O. Bernard**, and D. Friboulet. Analysis of motion tracking in echocardiographic image sequences : Influence of system geometry and point-spread function. *Ultrasonics*, 50(3) :373 – 386, 2010.
- [36] **O. Bernard**, D. Friboulet, P. Thevenaz, and M. Unser. Variational B-Spline Level-Set : A Linear Filtering Approach for Fast Deformable Model Evolution. *IEEE Transactions on Image Processing*, 18(6) :1179–1191, June 2009.

- [37] **O. Bernard**, B. Touil, J. D’hooge, and D. Friboulet. Statistical Modeling of the Radio-Frequency Signal for Partially- and Fully-Developed Speckle Based on a Generalized Gaussian Model with Application to Echocardiography. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 54(10) :2189–2194, October 2007.
- [38] A. Gelas, **O. Bernard**, D. Friboulet, and R. Prost. Compactly Supported Radial Basis Functions Based Collocation Method for Level-Set Evolution in Image Segmentation. *IEEE Transactions on Image Processing*, 16(7) :1873–1887, July 2007.
- [39] **O. Bernard**, J. D’hooge, and D. Friboulet. Statistics of the radio-frequency signal based on K distribution with application to echocardiography. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 53(9) :1689–1694, Sept 2006.
- [40] I. Dydenko, F. Jamal, **O. Bernard**, J. D’hooge, I. E. Magnin, and D. Friboulet. A level set framework with a shape and motion prior for segmentation and region tracking in echocardiography. *Medical Image Analysis*, 10(2) :162 – 177, 2006.

International conferences

I published **72** articles in international conferences.

- [41] S. Leclerc, E. Smistad, T. Grenier, C. Lartizien, A. Ostvik, F. Espinosa, P. M. Jodoin, L. Lovstakken, and **O. Bernard**. Deep learning applied to multi-structures segmentation in 2D echocardiography : a preliminary investigation of the required database size. In *2018 IEEE International Ultrasonics Symposium (IUS)*, pages 1–1, Oct 2018.
- [42] E. Smistad, A. Ostvik, I. Mjaland Salte, S. Leclerc, **O. Bernard**, and L. Lovstakken. Fully automatic real-time ejection fraction and MAPSE measurements in 2D echocardiography using deep neural networks. In *2018 IEEE International Ultrasonics Symposium (IUS)*, pages 1–1, Oct 2018.
- [43] F. Khellaf, S. Leclerc, J. D. Voorneveld, R. S. Bandaru, J. G. Bosch, and **O. Bernard**. Left ventricle segmentation in 3D ultrasound by combining structured random forests with active shape models. In *Proc. SPIE*, 2018.
- [44] S. Leclerc, T. Grenier, F. Espinosa, and **O. Bernard**. A fully automatic and multi-structural segmentation of the left ventricle and the myocardium on highly heterogeneous 2D echocardiographic data. In *2017 IEEE International Ultrasonics Symposium (IUS)*, pages 1–1, Sept 2017.
- [45] J. Pedrosa, S. Queirós, **O. Bernard**, J. Engvall, T. Edvardsen, E. Nagel, and J. D’hooge. Fast and fully automatic 3D left ventricular segmentation using shape-based B-spline explicit active surfaces. In *2017 IEEE International Ultrasonics Symposium (IUS)*, pages 1–1, Sept 2017.
- [46] A. Rodriguez-Molares, O. M. H. Rindal, **O. Bernard**, A. Nair, M. A. L. Bell, H. Liebgott, A. Austeng, and L. Lovstakken. The UltraSound ToolBox. In *2017 IEEE International Ultrasonics Symposium (IUS)*, pages 1–4, Sept 2017.
- [47] M. Zhang, H. Liebgott, F. Varray, D. Friboulet, and **O. Bernard**. Fourier-based ultrafast ultrasound imaging based on in-phase quadrature (IQ) data. In *2017 IEEE International Ultrasonics Symposium (IUS)*, pages 1–1, Sept 2017.

-
- [48] M. Zhang, F. Varray, L. Petrusca, D. Friboulet, H. Liebgott, and **O. Bernard**. Fourier-based 3D ultrafast ultrasound imaging with diverging waves : In vitro experiment validation. In *2017 IEEE International Ultrasonics Symposium (IUS)*, pages 1–1, Sept 2017.
- [49] Y. Zhou, M. De Craene, M. Sermesant, and **O. Bernard**. Phase-Based Registration of Cardiac Tagged MR Images by Incorporating Anatomical Constraints. In Tommaso Mansi, Kristin McLeod, Mihaela Pop, Kawal Rhode, Maxime Sermesant, and Alistair Young, editors, *Statistical Atlases and Computational Models of the Heart. Imaging and Modelling Challenges*, pages 39–47, Cham, 2017. Springer International Publishing.
- [50] H. Liebgott, A. Rodriguez-Molares, F. Cervenansky, J. A. Jensen, and **O. Bernard**. Plane-Wave Imaging Challenge in Medical Ultrasound. In *2016 IEEE International Ultrasonics Symposium (IUS)*, pages 1–4, Sept 2016.
- [51] M. Zhang, H. Liebgott, F. Varray, D. Friboulet, and **O. Bernard**. A Fourier-based formalism for 3D ultrafast imaging with diverging waves. In *2016 IEEE International Ultrasonics Symposium (IUS)*, pages 1–4, Sept 2016.
- [52] A. Besson, R. E. Carrillo, M. Zhang, D. Friboulet, **O. Bernard**, Y. Wiaux, and J. Thiran. Sparse regularization methods in ultrafast ultrasound imaging. In *2016 24th European Signal Processing Conference (EUSIPCO)*, pages 552–556, Aug 2016.
- [53] Y. Zhou, M. De Craene, and **O. Bernard**. Phase-based registration of cardiac tagged MR images using anatomical deformation model. In *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, pages 617–620, April 2016.
- [54] Y. Zhou, M. De Craene, O. Somphone, M. Sermesant, and **O. Bernard**. Generation of Realistic 4D Synthetic CSPAMM Tagged MR Sequences for Benchmarking Cardiac Motion Tracking Algorithms. In Sotirios A. Tsaftaris, Ali Gooya, Alejandro F. Frangi, and Jerry L. Prince, editors, *Simulation and Synthesis in Medical Imaging*, pages 108–117, Cham, 2016. Springer International Publishing.
- [55] R. E. Carrillo, A. Besson, M. Zhang, D. Friboulet, Y. Wiaux, J. Thiran, and **O. Bernard**. A Sparse regularization approach for ultrafast ultrasound imaging. In *2015 IEEE International Ultrasonics Symposium (IUS)*, pages 1–4, Oct 2015.
- [56] O. Lorintiu, H. Liebgot, **O. Bernard**, and D. Friboulet. Sub-sampled Doppler ultrasound reconstruction using block sparse Bayesian learning. In *2015 IEEE International Ultrasonics Symposium (IUS)*, pages 1–4, Oct 2015.
- [57] O. Lorintiu, H. Liebgott, A. Bernard, **O. Bernard**, and D. Friboulet. Compressed sensing reconstruction of line-wise sub-sampled 3D echographic images based on dictionary learning : an experimental study. In *2015 IEEE International Ultrasonics Symposium (IUS)*, pages 1–4, Oct 2015.
- [58] M. Zhang, A. Besson, R. E. Carrillo, F. Varray, M. Viallon, H. Liebgott, J. Thiran, D. Friboulet, and **O. Bernard**. Extension of Ultrasound Fourier Slice Imaging theory to sectorial acquisition. In *2015 IEEE International Ultrasonics Symposium (IUS)*, pages 1–4, Oct 2015.

- [59] M. Alessandrini, B. Heyde, S. Giffard-Roisin, H. Delingette, M. Sermesant, P. Allain, **O. Bernard**, M. De Craene, and J. D’hooge. Generation of ultra-realistic synthetic echocardiographic sequences to facilitate standardization of deformation imaging. In *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, pages 756–759, April 2015.
- [60] T. Dietenbeck, D. Barbosa, M. Alessandrini, J. D’hooge, D. Friboulet, and **O. Bernard**. A level-set approach for tracking objects in image sequences using a level conservation constraint : Application to cardiac sequences. In *2014 IEEE International Conference on Image Processing (ICIP)*, pages 863–866, Oct 2014.
- [61] O. Lortintiu, H. Liebgott, M. Alessandrini, **O. Bernard**, and D. Friboulet. Compressed sensing reconstruction of 3d ultrasound data using dictionary learning. In *2014 IEEE International Conference on Image Processing (ICIP)*, pages 1317–1321, Oct 2014.
- [62] N. Almeida, D. Barbosa, B. Heyde, R. O. Mada, D. Friboulet, **O. Bernard**, E. Samset, and J. D’hooge. Semi-automatic left-atrial segmentation from volumetric ultrasound using B-spline explicit active surfaces. In *2014 IEEE International Ultrasonics Symposium*, pages 612–615, Sept 2014.
- [63] **O. Bernard**, M. Zhang, F. Varray, P. Gueth, J. Thiran, H. Liebgott, and D. Friboulet. Ultrasound Fourier slice imaging : a novel approach for ultrafast imaging technique. In *2014 IEEE International Ultrasonics Symposium*, pages 129–132, Sept 2014.
- [64] M. Zhang, H. Liebgott, F. Varray, D. Friboulet, and **O. Bernard**. Speckle decorrelation of motion in Ultrasound Fourier images. In *2014 IEEE International Ultrasonics Symposium*, pages 1203–1206, Sept 2014.
- [65] M. De Craene, M. Alessandrini, P. Allain, S. Marchesseau, I. Waechter-Stehle, J. Weese, E. Saloux, H. G. Morales, R. Cuingnet, H. Delingette, M. Sermesant, **O. Bernard**, and J. D’hooge. Generation of ultra-realistic synthetic echocardiographic sequences. In *2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*, pages 73–76, April 2014.
- [66] D. Barbosa, N. Almeida, **O. Bernard**, D. Friboulet, E. Samset, and J. D’hooge. Segmentation of the left atrium of the heart using B-spline explicit active surfaces (BEAS). In *MICCAI-Challenge on Endocardial Three-dimensional Ultrasound Segmentation (CETUS)*, 2014.
- [67] A. Papachristidis, M. L. Geleijnse, E. Galli, B. Heyde, M. Alessandrini, D. Barbosa, M. J. Monaghan, **O. Bernard**, J. D’hooge, and J. G. Bosch. Clinical Expert Delineation of 3D Left Ventricular Echocardiograms for the CETUS Segmentation Challenge. In *MICCAI-Challenge on Endocardial Three-dimensional Ultrasound Segmentation (CETUS)*, 2014.
- [68] **O. Bernard**, B. Heyde, M. Alessandrini, D. Barbosa, S. Camarasu-Pop, F. Cervenansky, S. Valette, O. Mirea, E. Galli, M. Geleijnse, A. Papachristidis, J.G. Bosch, and J. D’hooge. Challenge on Endocardial Three-dimensional Ultrasound Segmentation (CETUS). In *MICCAI-Challenge on Endocardial Three-dimensional Ultrasound Segmentation (CETUS)*, 2014.

-
- [69] D. Barbosa, **O. Bernard**, B. Heyde, T. Dietenbeck, D. Friboulet, and J. D’hooge. Hybrid energy approach for real-time b-spline explicit active tracking of surfaces (heartBEATS). In *2013 IEEE International Ultrasonics Symposium (IUS)*, pages 663–666, July 2013.
- [70] D. Barbosa, **O. Bernard**, B. Heyde, T. Dietenbeck, D. Friboulet, and J. D’hooge. Towards online real-time strain estimation in volumetric us data : Feasibility study and initial clinical validation. In *2013 IEEE International Ultrasonics Symposium (IUS)*, pages 824–827, July 2013.
- [71] O. Lorintiu, H. Liebgott, **O. Bernard**, and D. Friboulet. Compressive sensing ultrasound imaging using overcomplete dictionaries. In *2013 IEEE International Ultrasonics Symposium (IUS)*, pages 45–48, July 2013.
- [72] T. Dietenbeck, D. Barbosa, M. Alessandrini, R. Jasaityte, V. Robesyn, J. D’hooge, D. Friboulet, and **O. Bernard**. Multiview myocardial tracking in echocardiographic 2D sequences using shape and motion constrained level-set. In *2013 IEEE 10th International Symposium on Biomedical Imaging*, pages 1022–1025, April 2013.
- [73] M. Alessandrini, A. Basarab, H. Liebgott, and **O. Bernard**. Cardiac motion assesement from echocardiographic image sequences by means of the structure multivector. In *2013 IEEE International Ultrasonics Symposium*, pages 1541–1544, 2013.
- [74] M. Alessandrini, H. Liebgott, D. Barbosa, and **O. Bernard**. Monogenic Phase Based Optical Flow Computation for Myocardial Motion Analysis in 3D Echocardiography. In Oscar Camara, Tommaso Mansi, Mihaela Pop, Kawal Rhode, Maxime Sermesant, and Alistair Young, editors, *Statistical Atlases and Computational Models of the Heart. Imaging and Modelling Challenges*, pages 159–168, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg.
- [75] D. Barbosa, B. Heyde, T. Dietenbeck, D. Friboulet, J. D’hooge, and **O. Bernard**. Fast Left Ventricle Tracking in 3D Echocardiographic Data Using Anatomical Affine Optical Flow. In S. Ourselin, D. Rueckert, and N. Smith, editors, *Functional Imaging and Modeling of the Heart*, pages 191–199, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg.
- [76] S. Queirós, D. Barbosa, B. Heyde, P. Morais, D. Friboulet, P. Claus, **O. Bernard**, and J. D’hooge. Fast Fully Automatic Segmentation of the Myocardium in 2D Cine MR Images. In S. Ourselin, D. Rueckert, and N. Smith, editors, *Functional Imaging and Modeling of the Heart*, pages 71–79, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg.
- [77] M. Alessandrini, H. Liebgott, D. Friboulet, and **O. Bernard**. Simulation of realistic echocardiographic sequences for ground-truth validation of motion estimation. In *2012 19th IEEE International Conference on Image Processing*, pages 2329–2332, Sept 2012.
- [78] F. Galluzzo, N. Speciale, and **O. Bernard**. A rigorous and efficient GPU implementation of level-set sparse field algorithm. In *2012 19th IEEE International Conference on Image Processing*, pages 1705–1708, Sept 2012.
- [79] M. Alessandrini, H. Liebgott, A. Basarab, P. Clarysse, and **O. Bernard**. Monogenic signal for cardiac motion analysis from tagged magnetic resonance image sequences. In *Computing in Cardiology*, pages 685–688, 2012.

- [80] M. Alessandrini, H. Liebgott, D. Friboulet, and **O. Bernard**. Monogenic phase based myocardium motion analysis from cardiac ultrasound with transverse oscillations. In *2012 IEEE International Ultrasonics Symposium*, pages 1098–1101, 2012.
- [81] D. Barbosa, T. Dietenbeck, D. Friboulet, J. D’hooge, and **O. Bernard**. Fast and fully automatic 3D echocardiographic segmentation using B-spline explicit active surfaces. In *2012 IEEE International Symposium on Biomedical Imaging*, pages 1088–1091, 2012.
- [82] D. Barbosa, **O. Bernard**, B. Heyde, T. Dietenbeck, H. Houle, D. Friboulet, and J. D’hooge. B-spline Explicit Active Tracking of Surfaces (BEATS) : Application to real-time 3D segmentation and tracking of the left ventricle in 3D echocardiography. In *2012 IEEE International Ultrasonics Symposium*, pages 224–227, 2012.
- [83] D. Bibicu, D. Barbosa, B. Heyde, D. Friboulet, **O. Bernard**, L. Moraru, P. Claus, and J. D’hooge. An integrated solution for semi-automatic segmentation of volumetric ultrasound data based on B-spline explicit active surfaces. In *2012 IEEE International Ultrasonics Symposium*, pages 2643–2646, 2012.
- [84] L. Flórez-Valencia, E. E. Dávila Serrano, J. G. Riveros Reyes, **O. Bernard**, J. Latt, O. Malaspinas, B. Chopard, G. Courbebaisse, and M. Orkisz. Virtual deployment of pipeline flow diverters in cerebral vessels with aneurysms to understand thrombosis. In *MICCAI-Workshop on Computer Assisted Stenting*, pages 49–56, 2012.
- [85] F. Galluzzo, D. Barbosa, H. Houle, N. Speciale, D. Friboulet, J. D’hooge, and **O. Bernard**. A GPU Level-Set Segmentation Framework for 3D Echocardiography. In *2012 IEEE International Ultrasonics Symposium*, pages 2639–2642, 2012.
- [86] D. Barbosa, **O. Bernard**, T. Dietenbeck, B. Heyde, D. Friboulet, H. Houle, and J. D’hooge. Fast 3D echocardiographic segmentation using B-Spline Explicit Active Surfaces : A validation study in a clinical setting. In *2011 IEEE International Ultrasonics Symposium*, pages 1854–1857, Oct 2011.
- [87] F. Douziech, H. Liebgott, **O. Bernard**, and A. Basarab. Affine phase based motion estimation applied to echocardiography. In *2011 IEEE International Ultrasonics Symposium*, pages 736–739, Oct 2011.
- [88] D. Barbosa, J. D’hooge, T. Dietenbeck, D. Friboulet, and **O. Bernard**. Towards real-time 3D region-based segmentation : B-spline explicit active surfaces. In *2011 18th IEEE International Conference on Image Processing*, pages 3121–3124, Sept 2011.
- [89] D. Barbosa, T. Dietenbeck, D. Friboulet, J. D’hooge, and **O. Bernard**. Real-time region-based segmentation of 3D inhomogeneous objects in medical images. In *2011 IEEE International Symposium on Biomedical Imaging : From Nano to Macro*, pages 1986–1989, March 2011.
- [90] T. Dietenbeck, M. Alessandrini, D. Barbosa, J. D’hooge, D. Friboulet, and **O. Bernard**. Multiview myocardial segmentation in echocardiographic images using a piecewise parametric shape prior. In *2011 IEEE International Symposium on Biomedical Imaging : From Nano to Macro*, pages 645–648, March 2011.
- [91] S. Maggio, M. Alessandrini, N. Speciale, **O. Bernard**, D. Vray, O. Basset, and M. Unser. Continuous-domain ARMA modeling for ultrasound tissue characterization. In *2011 IEEE*

-
- International Symposium on Biomedical Imaging : From Nano to Macro*, pages 895–898, March 2011.
- [92] M. Alessandrini, S. Maggio, J. Porée, L. De Marchi, N. Speciale, E. Franceschini, **O. Bernard**, and O. Basset. An expectation maximization framework for an improved tissue characterization using ultrasounds. In *SPIE Medical Imaging*, 2011.
- [93] T. Dietenbeck, M. Alessandrini, D. Friboulet, and **O. Bernard**. Creaseg : A free software for the evaluation of image segmentation algorithms based on level-set. In *2010 IEEE International Conference on Image Processing*, pages 665–668, Sept 2010.
- [94] M. Alessandrini, T. Dietenbeck, D. Barbosa, J. D’hooge, O. Basset, N. Speciale, D. Friboulet, and **O. Bernard**. Segmentation of the Full Myocardium in Echocardiography Using Constrained Level-Sets. In *Computing in Cardiology*, pages 409–412, 2010.
- [95] M. Alessandrini, T. Dietenbeck, O. Basset, D. Friboulet, and **O. Bernard**. Using a geometric formulation of annular-like shape priors for constraining variational level-sets. In *IEEE International Conference on Image Processing*, pages 669–672, 2010.
- [96] D. Barbosa, **O. Bernard**, O. Savu, T. Dietenbeck, B. Heyde, P. Claus, D. Friboulet, and J. D’hooge. Coupled B-spline Active Geometric Functions for Myocardial Segmentation in Cardiac Ultrasound. In *2010 IEEE International Ultrasonics Symposium*, pages 1648–1651, 2010.
- [97] J. Vandemeulebroucke, **O. Bernard**, J. Kybic, P. Clarysse, and D. Sarrut. Automatic motion mask extraction for deformable registration of the lungs. In *XVIIth International Conference on the Use of Computers in Radiation Therapy*, 2010.
- [98] M. Alessandrini, D. Friboulet, O. Basset, J. D’hooge, and **O. Bernard**. Level-set segmentation of myocardium and epicardium in ultrasound images using localized Bhattacharyya distance. In *2009 IEEE International Ultrasonics Symposium*, pages 2468–2471, Sept 2009.
- [99] **O. Bernard** and D. Friboulet. Fast medical image segmentation through an approximation of narrow-band B-spline level-set and multiresolution. In *2009 IEEE International Symposium on Biomedical Imaging : From Nano to Macro*, pages 45–48, June 2009.
- [100] B. Touil, A. Basarab, **O. Bernard**, and D. Friboulet. Influence of system geometry on motion tracking in echocardiographic image sequences. In *2009 IEEE International Symposium on Biomedical Imaging : From Nano to Macro*, pages 1150–1153, June 2009.
- [101] H. Liebgott, A. Basarab, S. Marincas, **O. Bernard**, and D. Friboulet. Tangential oscillations for motion estimation in echocardiography. In *2008 IEEE Ultrasonics Symposium*, pages 1761–1764, Nov 2008.
- [102] L. Saroul, **O. Bernard**, D. Vray, and D. Friboulet. Prostate segmentation in echographic images : A variational approach using deformable super-ellipse and rayleigh distribution. In *2008 5th IEEE International Symposium on Biomedical Imaging : From Nano to Macro*, pages 129–132, May 2008.

- [103] **O. Bernard**, D. Friboulet, P. Thevenaz, and M. Unser. Variational B-spline level-set method for fast image segmentation. In *2008 5th IEEE International Symposium on Biomedical Imaging : From Nano to Macro*, pages 177–180, May 2008.
- [104] B. Touil, **O. Bernard**, and D. Friboulet. Motion decorrelation in echocardiography : Analysis from a realistic simulation. In *2008 5th IEEE International Symposium on Biomedical Imaging : From Nano to Macro*, pages 1469–1472, May 2008.
- [105] S. Camarasu-Pop, H. Benoit-Cattin, L. Guigues, P. Clarysse, **O. Bernard**, and D. Friboulet. Towards a Virtual Radiological Platform Based on a Grid Infrastructure. In *Medical imaging on grids : achievements and perspectives (MICCAI Grid Workshop)*, pages 85–95, 2008.
- [106] H. Liebgott, **O. Bernard**, C. Cachard, and D. Friboulet. Field Simulation Parameters Design for Realistic Statistical Parameters of Radio - Frequency Ultrasound Images. In *2007 IEEE Ultrasonics Symposium Proceedings*, pages 2247–2250, Oct 2007.
- [107] A. Gelas, J. Schaerer, **O. Bernard**, D. Friboulet, P. Clarysse, I. E. Magnin, and R. Prost. Radial Basis Functions Collocation Methods for Model Based Level-Set Segmentation. In *2007 IEEE International Conference on Image Processing*, volume 2, pages II – 237–II – 240, Sept 2007.
- [108] **O. Bernard**, B. Touil, A. Gelas, R. Prost, and D. Friboulet. A RBF-Based Multiphase Level Set Method for Segmentation in Echocardiography using the Statistics of the Radiofrequency Signal. In *2007 IEEE International Conference on Image Processing*, volume 3, pages III – 157–III – 160, Sept 2007.
- [109] **O. Bernard**, B. Touil, A. Gelas, R. Prost, and D. Friboulet. Segmentation of Myocardial Regions in Echocardiography Using the Statistics of the Radio-Frequency Signal. In Frank B. Sachse and Gunnar Seemann, editors, *Functional Imaging and Modeling of the Heart*, pages 433–442, Berlin, Heidelberg, 2007. Springer Berlin Heidelberg.
- [110] **O. Bernard**, J. D’hooge, and D. Friboulet. Statistical Modeling of the Radio-Frequency Signal in Echocardiographic Images Using Modified K-Distribution and Generalized Generalized Gaussian Distribution. In *2006 IEEE Ultrasonics Symposium*, pages 1710–1713, Oct 2006.
- [111] **O. Bernard**, J. D’Hooge, and D. Friboulet. Statistical modeling of the radio-frequency signal in echocardiographic images based on generalized Gaussian distribution. In *3rd IEEE International Symposium on Biomedical Imaging : Nano to Macro, 2006.*, pages 153–156, April 2006.
- [112] **O. Bernard**, J. D’hooge, and D. Friboulet. Segmentation of echocardiographic images based on statistical modelling of the radio-frequency signal. In *2006 14th European Signal Processing Conference*, pages 1–5, 2006.

Invited talks

I participated in 5 invited talks.

- [113] M. Alessandrini, A. Basarab, M. De Craene, M. Sermesant, H. Liebgott, **O. Bernard**, and J. D’hooge. The role of the image phase in cardiac strain imaging. In *2015 23rd European Signal Processing Conference (EUSIPCO)*, pages 2796–2800, Aug 2015.
- [114] **O. Bernard**, D. Barbosa, M. Alessandrini, D. Friboulet, and J. D’hooge. 3D Echocardiography : recent advances and future directions. In *VipIMAGE2013 - ECCOMAS Thematic Conferences on Computational Vision and Medical Image Processing*, Oct 2013.
- [115] A. Basarab, H. Liebgott, **O. Bernard**, D. Friboulet, and D. Kouamé. Medical ultrasound image reconstruction using distributed compressive sampling. In *2013 IEEE 10th International Symposium on Biomedical Imaging*, pages 628–631, April 2013.
- [116] H. Liebgott, A. Basarab, D. Kouamé, **O. Bernard**, and D. Friboulet. Compressive sensing in medical ultrasound. In *2012 IEEE International Ultrasonics Symposium*, pages 1–6, 2012.
- [117] D. Friboulet and **O. Bernard**. Segmentation of medical ultrasound images. In *2011 IEEE International Ultrasonics Symposium*, 2011.

Book chapters

I participated in the writing of **2** book chapters.

- [118] **O. Bernard**, P. Clarysse, T. Dietenbeck, D. Friboulet, S. Jehan-Besson, and J. Pousin. *Extraction and Segmentation of Structures in Image Sequences*, chapter 1, pages 1–64. Wiley-Blackwell, 2015.
- [119] M. Alessandrini, A. Basarab, **O. Bernard**, and P. Delachartre. *Estimation du mouvement du coeur basée sur la phase en imagerie cardiaque multimodalités*, pages 221–248. Hermès-Lavoisier, 2014.

Research activities

General introduction

During my PhD, I had the opportunity to work on the segmentation of echocardiographic images with a particular interest in the statistics of the ultrasonic signal in order to improve the efficiency of the data attachment term involved in a dedicated active contour model. The main innovations developed during my PhD thus concern a complete study on the statistics of the ultrasonics radiofrequency signal and the derivation of a collocation method based on radial basis functions to efficiently model level-sets evolution in image segmentation. This work was valued by the publication of 4 articles (2 as first author) in international journals and 6 articles (5 as first author) in international conferences.

The research activities I have carried out since then are described in the following chapters. My work concerns signal and image processing applied to medical imaging. In particular, I focused my intention on three main topics : cardiac segmentation, cardiac motion estimation and ultrasound image reconstruction, all for a better diagnosis of cardiac diseases. In terms of methodological developments, I worked on *i*) variational approaches through level-set and B-spline explicit active surface formalisms; *ii*) statistical learning through random forest and deep learning techniques; *iii*) signal modeling based on the monogenic signal theory; *iv*) inverse problem based on the Fourier slice theorem; *v*) optimization problem based on personalized anatomical regularization. In terms of application, I dedicated my research on cardiac imaging (mainly ultrasound and MRI), both in terms of image reconstruction and image analysis. The clinical goal of my research concerns a better diagnosis of heart diseases through reliable and robust measurements from the image of clinical indices such as ventricular volumes, ejection fraction and myocardial strain and strain rate.

In terms of fundings, I financed my research through four main sources :

- national fundings thanks to ministerial scholarship (co-supervision of 1 PhD), LABEX entity (co-supervision of 2 PhDs and 1 postdoc with the LABEX PRIMES) and ANR project (co-supervision of 1 postdoc from the US-tagging ANR project of professor Hervé Liebgott) ;
- European fundings (co-supervision of 2 PhDs from the ERC grant of professor Jan D’hooge) ;
- international fundings (co-supervision of 1 PhD from the China Scholarship Council) ;
- industrial fundings thanks to the French CIFRE doctoral scholarships (2 PhDs co-supervised with Philips Medisys society).

Cardiac segmentation : toward robust volumetric indices estimation

1 Introduction

Analysis of cardiac function plays an important role in clinical cardiology for patient management, disease diagnosis, risk evaluation, and therapy decision [80, 178, 283]. Thanks to digital imagery, the assessment of a set of complementary indices computed from different structures of the heart is a routine task for cardiac diagnostics.

Because of its well-known capacity for discriminating different types of tissues, Cardiac Magnetic Resonance Imaging (CMR) (built from series of parallel short axis slices) is considered as the gold standard of cardiac function analysis through the assessment of the left and right ventricular ejection fractions (EF) and stroke volumes (SV), the left ventricle (LV) mass and the myocardium thickness. This requires accurate delineation of the left ventricular endocardium and epicardium, and of the right ventricular endocardium for both end diastolic (ED) and end systolic (ES) phase instances. In clinical practice, semi-automatic segmentation is still a daily practice because of the lack of accuracy of fully-automatic cardiac segmentation methods. This leads to time consuming tasks prone to intra- and inter-observer variability [164]. The difficulties of CMR segmentation have been clearly identified [213] : *i*) presence of poor contrast between myocardium and surrounding structures (conversely, there is a high contrast between blood and the myocardium) ; *ii*) brightness heterogeneities in the left ventricular/right ventricular cavities due to blood flow ; *iii*) presence of trabeculae and papillary muscles with intensities similar to the myocardium ; *iv*) non-homogeneous partial volume effects due to the limited CMR resolution along the long-axis ; *v*) inherent noise due to motion artifacts and heart dynamics ; *vi*) shape and intensity variability of the heart structures across patients and pathologies ; *vii*) presence of banding artifact.

While CMR remains the gold standard for the assessment of cardiac morphology and function, real-time 2D echocardiography (RT2DE) is the mostly used modality to diagnose heart pathology in clinical routine. This attractiveness may be explained by its bedside applicability, excellent temporal resolution, real-time character, low cost and absence of ionizing radiation. Unfortunately, as for CMR, semi-automatic or manual annotation from ultrasound images is still daily work in clinical routine due to the lack of accuracy and reproductibility of fully-automatic cardiac segmentation methods. This also leads to time consuming tasks prone to intra- and inter-observer variability [13]. The inherent difficulties for segmenting echocardiographic images have been well documented : *i*) poor contrast between heart tissues and the blood pool ; *ii*) brightness inhomogeneities ; *iii*) variation in the speckle pattern along the myocardium due to the orientation of the cardiac probe with respect to tissue ; *iv*) presence of trabeculae and papillary

muscles with intensities similar to the myocardium ; *v*) significant tissue echogenicity variability within the population ; *vi*) shape, intensity and motion variability of the heart structures across patients and pathologies ; *vii*) out-of-plane motion.

Over the last decade the assessment of cardiac morphology and function by ultrasound imaging has made a significant step forward by the introduction of real-time 3D echocardiography (RT3DE), as it allows a truly 3D visualization of the heart avoiding some of the problems intrinsically associated with 2D imaging such as foreshortening, out-of-plane motion and the need of geometric assumptions for volume estimation [169]. Unfortunately, due to the intrinsic physical limits of acoustical wave propagation, RT3DE currently suffers from a low spatial and temporal resolution compared to conventional 2D echocardiography and the potential presence of motion artifacts due to stitching strategies [136]. As a consequence, state-of-the-art commercial solutions towards LV segmentation still require some degree of user interaction both at the initialization step and after segmentation/tracking when corrections are required [172]. The development of fully automatic and fast techniques for LV volumetric assessment is therefore still an open issue and therefore an active field of research [147]. The majority of methods that estimate cardiac volumes go through a segmentation step of the anatomical structures of interest [26, 28, 36, 67, 111, 126, 183, 197, 272].

The remainder of the chapter is organized as follows.

- Sec. 2 : Review of state-of-the-art methods in cardiac image segmentation ;
- Sec. 3 : Presentation of the different datasets I have set up to assess the performance of segmentation and clinical indices extraction methods ;
- Sec. 4 : Detailed description of a novel formalism, named B-spline Explicit Active Surface (BEAS), that we introduced to efficiently segment real-time 3D echocardiography ;
- Sec. 5 : Presentation of a study we put in place to evaluate how far the state-of-the-art deep convolutional network methods can go at assessing 2D echocardiographic images ;
- Sec. 6 : Overview of a study that we realized to measure how far state-of-the-art deep learning methods can go at segmentation cardiac structures as well as classifying pathologies.
- Sec. 7 : Conclusions of this chapter

2 State-of-the-art in cardiac image segmentation

2.1 Cardiac ultrasound

Several surveys of echocardiographic segmentation methods have been proposed, both in 2D [41, 177] and 3D [30, 145]. Most of the reported methods focused on the segmentation of the LV_{Endo} border. Techniques successfully applied for LV segmentation can be categorized based on their underlying methodology : deformable models, surface fitting approaches, graph-cuts, multi-atlas and machine learning techniques. Deformable models constitute the most widely used technique to segment echocardiography data because of its attractiveness in terms of mathematical formulation of the segmentation problem, its flexibility in terms of shape representation and its computational time efficiency (depending on the chosen representation) [11, 23, 53, 69]. Surface fitting approaches are based on shape interpolation from a set of points that were previously identified as belonging to the endocardial surface. The most popular techniques that have been applied in echocardiography are active shape models (ASM) [273] and Doo-Sabin subdivision surfaces (DSS) [185, 222]. The main feature of ASM corresponds to the embedded representation of the shape based on a space built from a set of references.

These methods are known to require a large number of training samples to be meaningful. The main asset of DSS is its capacity to represent smooth surfaces by a small number of parameters. Graph-cuts and multi-atlas techniques have also been applied in echocardiography [124, 303]. In graph-cut techniques, the underlying graph is usually expressed in a coordinate system adapted to the anatomy of the endocardial surface. The attractiveness of such approach resides in its capacity of expressing the segmentation problem as a graph and thus benefits from powerful graph-flow optimization techniques. Multi-atlas segmentation techniques are registration-based techniques that perform the segmentation of the endocardium from a set of pre-segmented volumes. Based on the selection of the best atlases after registration, merging techniques are usually used to enhance the quality of the segmented results. Finally, standard machine learning approaches have proven their value for LV segmentation in several studies [141, 222, 288]. Usually, these techniques are not designed to produce the final segmented surface but provide a fundamental intermediate step that guides the segmentation process. For example, methods based on marginal space learning have been specially designed to accurately locate the myocardial region [288]. Similarly, a boundary fragment model has been successfully applied for the identification of extracted contours [222]. Finally, random forest-based approaches have been applied as a discriminative classifier in order to characterize the affiliation of each voxel to the myocardium [141]. Table 1 gives a summary of the accuracy of existing automatic 3D LV segmentation techniques when compared with manual references. Although this table provides some evidence of the relative performance of the different approaches, it is clear that a fair and true comparison is not feasible given the differences in patient population and image quality (due to different acquisition conditions and equipment). This points to the need for a publicly available common database.

TABLE 1 – State-of-the-art algorithms for LV segmentation in RT3DE.

Study *	Algorithm		Volumes						Segmented Shapes		
	#	T	R			BA ($\mu \pm 2\sigma$)			MAD ($\mu \pm \sigma$)		
			EDV	ESV	EF	EDV (ml.)	ESV (ml.)	EF (%)	ED (mm)	ES (mm)	FC (mm)
Angelini <i>et al.</i> [11]	10		0.63	0.62	0.45	16.1±50	6.6±34	0.5±22			
Hansegard <i>et al.</i> [100]	21	0.04	0.91	0.91	0.74	-5.9±21	6.2±19	-7.7±12			2.2±0.6
Leung <i>et al.</i> [144]	99		0.95			-1.47±40			2.9±1.0		
Duan <i>et al.</i> [71]	35	0.033									4.0±3.2
Leung <i>et al.</i> [143]	35	6	0.982			1.9±14					1.35±0.5
Yang <i>et al.</i> [288]	67	1.5				1.32±12					1.28±1.1
Rajpoot <i>et al.</i> [218]	34					-5.0±49	1.2±26	-0.7±14	2.2±0.7	1.52±0.4	
Dikici <i>et al.</i> [65]	29	0.08							2.0±X	2.0±X	
Barbosa <i>et al.</i> [23]	24	1	0.97	0.97	0.91	-1.4±23	2±19	-1.0±10			
Pedrosa <i>et al.</i> [197]	30	0.8	0.95	0.96	0.91	-3.3±19	-4.8±16	1.7±5	1.8±0.6	2.0±0.7	
Oktay <i>et al.</i> [183]	30				0.91			1.8±10	1.9±0.5	2.1±0.8	

* The following symbols were used : EDV : End Diastolic Volume ; ESV : End Systolic Volume ; EF : Ejection Fraction ;

* FC : full cycle ; # : number of exams ; T : average frame processing time (s) per volume ; R : correlation coefficient ;

BA : Bland-Altman analysis ; MAD : mean absolute surface distance.

Deep-learning methods have been successfully applied to the segmentation of the LV_{Endo} in echocardiography. In 2012, Carneiro *et al.* developed a two-stage deep learning method for the segmentation of the LV_{Endo} for 2D echocardiographic images restricted to four-chambers view acquisitions [41]. Based on a maximum *a posteriori* framework, the authors formulated the LV segmentation problem according to two successive steps : *i*) the automatic selection of several regions in the tested image where the LV_{Endo} is fully present ; *ii*) the automatic extraction of the LV_{Endo} contour from the previously selected regions. These two steps involve a deep belief

network. Their method was trained on 400 images from 12 sequences of different patients with various pathologies and tested on 50 images from 2 sequences of healthy subjects. They obtained an average Hausdorff distance of ~ 18 mm and an average mean absolute distance of ~ 8 mm for the LV_{Endo} . In 2017, Smistad *et al.* [246] showed that the U-Net CNN [225] could be trained to successfully segment the left ventricle in 2D ultrasound images. Due to lack of training data, the network was trained with the output of a state-of-the-art deformable model segmentation method [244]. On a manually segmented test set, the results showed that the network and the deformable model obtained the same accuracy with a Dice score of 0.87. Recently, Oktay *et al.* [183] used CNNs to segment the 3D LV_{Endo} structure. The core of their neural network is based on an architecture similar to the 3D U-Net [179], whose segmentation output is constrained to fit a non-linear compact representation of the underlying anatomy derived from an auto-encoder network. The performance of their method was assessed on the CETUS database [30]. They obtained the following scores for the segmentation of the 3D LV_{Endo} structure : *i)* average Dice values of 0.912 (ED) and 0.873 (ES) ; *ii)* average Hausdorff distances of 7.0 mm (ED) and 7.7 mm (ES) and *iii)* average mean absolute distances of 1.9 mm (ED) and 2.1 mm (ES) [183]. Interestingly, the use of only 15 patients during the training phase illustrates the strong potential of deep learning techniques to analyze echocardiographic images.

2.2 Cardiac MRI

Petitjean *et al.* proposed in 2011 a complete review of segmentation methods for delineating the left ventricle (LV) and/or the right ventricle (RV) in short axis cardiac MR images [200]. In this study, the authors listed the results published in more than 70 peer-reviewed publications. The reported methods can be divided in two main categories : weak prior and strong prior methods. The first group involves weak assumptions such as spatial, intensity or anatomical information. It includes image-based techniques (threshold, dynamic programming) [148], pixel classification methods (clustering, Gaussian mixture model fitting) [269], deformable models (active contour, level-set) [45] and graph-based approaches (graph-cut) [18]. The second group uses methods with strong prior including shape prior based deformable models [213], active shape and appearance models [165] and atlas based methods [19], all requiring a training dataset with manual annotations. Although this huge work provides a complete picture of the performance of the state-of-the-art methods in LV/RV segmentation, it does not benchmark these techniques with a unique dataset. Such comparison thus remains a glaring issue in our community.

To our knowledge, before 2013 no deep learning techniques was used to analyze CMRI. However, a drastic change occurred in 2015 during the Kaggle Second Annual Data Science Bowl during which the undeniable power of deep learning methods was revealed to the community. Since then, a dozen deep learning papers have been published on the topic of CMRI analysis. Most papers used 2D convolutional neural networks (CNNs) and analyzed the MRI data slice by slice. Three papers used deep learning framework to extract relevant features for segmentation. Emad *et al.* [81] used a patch-wise CNN to localize the LV in CMRI slices. Kong *et al.* [130] developed a temporal regression framework to identify end-diastolic and end-systolic instances from the cardiac cycle by integrating a 2D CNN with a recurrent neural network (RNN). The CNN was used to encode the spatial information while the RNN was used to decode the temporal information. Finally Zhang *et al.* [294] used a simple CNN to automatically detect missing slices (apical and basal) in cardiac exams to assess the quality of MRI acquisitions. Four papers used deep learning methods combined with classical cardiac segmentation tools. Rupprecht *et al.* [228] integrated a patch-based CNN into a semi-automatic active contour (a snake) to segment

cardiac structures. Ngo *et al.* [175] used a deep belief network (DBN) to accurately initialize and guide a level-set model to segment the left ventricle. Yang *et al.* [287] developed a combined approach between CNN and multi-atlas to perform LV segmentation. In particular, a deep architecture was trained to learn deep features achieving optimal performance for the label fusion operation classically involved in multi-atlas segmentation. Alternatively, Avendi *et al.* [16] proposed a combined deep-learning and deformable-model approach to automatically segment the left ventricle. The method works as follows : *i*) a simple CNN locates and crops the LV ; *ii*) a stack of autoencoders pre-segment the LV shape ; *iii*) the pre-segmented shape is refined with a deformable model. Although the authors report almost perfect results on Sunnybrook 2009, it is not clear how their method generalizes to more than one cardiac region. Finally, three papers used standalone deep learning techniques to segment cardiac structures from CMR data. Poudel *et al.* [204] proposed a recurrent fully-convolutional network (RFCN) that learns image representations from the full stack of 2D slices. The derived architecture allows leveraging inter-slice spatial dependences through internal memory units. Tran *et al.* [265] developed a deep fully convolutional neural network architecture to segment both LV and RV structures. Finally, Oktay *et al.* [182] proposed an image super-resolution approach based on a residual convolutional neural network model. Their key idea is to reconstruct high resolution 3D volumes from 2D image stacks for more accurate image analysis. For more details on deep learning methods applied to medical image analysis (including cardiac MRI segmentation) please refer to Litjens *et al.* [147] and Havaei *et al.* [102].

2.3 Previous cardiac datasets

To our knowledge, before the two projects we set up from 2014 (see Section 3), no publicly-available ultrasound datasets was provided in the community to assess the performance of segmentation and clinical indices extraction methods. With regard to MRI, four large datasets of clinical data have been broadly accepted by the community in the last decade. These datasets were released in conjunction with an international challenge allowing the organizers to benchmark state-of-the-art methods.

The Sunnybrook Cardiac MR Left Ventricle Segmentation challenge - MICCAI 2009² provides a database of 45 cardiac cine-MR images from four different pathological groups namely : heart failure with ischemia, heart failure without ischemia, hypertrophic cardiomyopathy, and normal subjects. The data is provided with two manually-drawn contours, one for the endocardium and one for the epicardium [215]. Although the database is still publicly available, neither collated results nor comparative study have been published thus reducing the impact of this event. However, recent papers [16, 213, 253] reported results from several automatic and semi-automatic segmentation methods published since the 2009 challenge. According to those results, the top performing methods (many of which being only focused on the endocardium segmentation) report Dice scores between 0.90 and 0.94 for the endocardium and/or the epicardium and an average perpendicular distance of less than 2.0 *mm* and an average 2D Hausdorff distance between 3.0 and 5.0 *mm*.

The LV Segmentation Dataset and Challenge, MICCAI-STACOM 2011³ focuses on the comparison of LV segmentation methods [251]. The database is made of CMR acquisitions from 200 patients with coronary artery disease and prior myocardial infarction (100

2. http://smial.sri.utoronto.ca/LV_Challenge/Home.html

3. www.cardiacatlas.org/challenges/lv-segmentation-challenge/

for training and 100 for testing). In this study, the authors introduced the concept of objective ground truth based on the evidence from the contribution of several raters. In particular, ground truths computed for the 100 patients of the testing set were generated from an Expectation-Maximization framework (the STAPLE algorithm) [282] using the results of two fully-automated methods (automated raters) and three semi-automated approaches with manual input (manual raters). No 100% manually annotated ground truth were involved in this study. From the derived ground truths, the best results in terms of segmentation accuracy were obtained by a guide-point modeling technique (manual rater) which obtained an average Jaccard score of 0.84 [146].

The Right Ventricle Segmentation Dataset - MICCAI 2012⁴ aims at comparing RV segmentation methods based on a set of 48 cardiac cine-MR data with contours drawn by one cardiac radiologist (16 for training, 32 for testing) [201]. Three fully-automatic and four semi-automatic methods were evaluated through this challenge. Back in 2012, the outcome of the challenge revealed that the best scores were obtained by semi-automatic methods like the graph-cut method by Grosgeorge *et al.* [95] which reached an average Dice score of 0.78 and an average 2D Hausdorff distance of 8.62. In a recent publication, Phi Vu Tran [265] showed how a fine-tuned fully-convolutional neural network [151] can out-perform every semi-automatic method with an average Dice score of 0.85.

The 2015 Kaggle Second Annual Data Science Bowl⁵ is a challenge for which more than 190 teams competed to win the \$200,000 grand prize. The goal of this event was to automatically measure ED and ES volumes from CMR. Challengers were given a database composed of 500 patients for training and 200 patients for testing. The training images came only with the ED and ES reference volumes and not a manually segmented ground truth as for the other three datasets. The outcome of the challenge revealed that the top-performing methods relied on deep learning technologies, in particular fully convolutional networks (fCNN) [151] and U-Net [225]. Unfortunately, no summary paper was provided in the wake of this challenge.

Table 2 summaries the MRI cardiac datasets mentioned above. Let us also mention that other fully-annotated cardiac datasets have been released such as HVSMR 2016⁶ and the Multi-Modality Whole Heart Segmentation dataset⁷. Although interesting, these datasets contain images that are clinically atypical. Furthermore, without being bound to a challenge, the UK Biobank [199] corresponds to the largest existing CMR database which could be used to train and test deep learning methods whenever the manual annotations of these images will be rendered public. However, one limit of this database is that it is not free, which inevitably limits its access by research teams, and thus does not correspond to open science initiatives such as challenges.

3 1st contribution : setting up open-access datasets

In order to get the possibility to assess the potential of state-of-the-art methods in segmenting cardiac images and extracting relevant clinical indices, we spent lots of time to design several open-access datasets with manual annotations from cardiologists and radiologist experts.

4. www.litislab.fr/?projet=1rvsc

5. www.kaggle.com/c/second-annual-data-science-bowl

6. <http://segchd.csail.mit.edu/>

7. <http://stacom2017.cardiacatlas.org/>

TABLE 2 – Summary of the full set of existing cardiac MRI datasets which are publicly available for comparison purposes.

CMRI datasets								
Name	Year	Nb Subjects		Ground truth				Active website
		train	test	LV	RV	Myo	Pathology	
Sunnybrook	2009	45	—	✓	✗	✓	✓	✗
STACOM	2011	100	100	✓	✗	✓	✗	✗
MICCAI RV	2012	16	32	✗	✓	✗	✗	✗
Kaggle	2015	500	200	✗	✗	✗	✗	✗
ACDC	2017	100	50	✓	✓	✓	✓	✓

3.1 CAMUS dataset

This large scale dataset has been set-up to evaluate how far state-of-the-art deep convolutional networks and other non-deep learning methods can go at assessing 2D echocardiographic images, *i.e.* segmenting cardiac structures as well as estimating clinical indices. The full dataset is made available for download at <https://camus.creatis.insa-lyon.fr/challenge/>.

3.1.1 Patient selection

The proposed dataset consists of clinical exams from 500 patients, acquired at the University Hospital of St Etienne (France) and included in this study within the regulation set by the local ethical committee of the hospital. The acquisitions were optimized to perform LV_{EF} measurements. In order to enforce clinical realism, neither prerequisite nor data selection have been performed. Consequently, *i)* some cases were difficult to trace; *ii)* the dataset involves a wide variability of acquisition settings; *iii)* for some patients, parts of the wall were not visible in the images; *iv)* for some cases, the probe orientation recommendation to acquire a rigorous four-chambers view was simply impossible to follow and a five-chambers view was acquired instead. This produced a highly heterogeneous dataset, both in terms of image quality and pathological cases, which is typical of daily clinical practice data. Table 3 provides the main information which characterizes the collected dataset. From this table, one can see that half of the dataset population has a LV_{EF} lower than 45%, thus being considered at pathological risk (beyond the uncertainty of the measurement). Also, 19% of the images have a poor quality (based on the opinion of one expert O_{1a}), indicating that for this subgroup the localization of the LV_{Endo} and LV_{Epi} as well as the estimation of clinical indices are not considered clinically accurate and workable. In classical analysis, poor quality images are usually removed from the dataset because of their clinical uselessness. Therefore, those data were not involved in this project during the computation of the different metrics but were used to study their influence as part of the training and validation sets for deep learning techniques. The dataset was divided into 10 folds to perform standard cross-validation for the machine learning methods. Each fold contains 50 patients with the same distributions in terms of image quality and LV_{EF} as the full dataset (see table 3). For each of the 10 test sets, the remaining 450 patients (9 folds) were used during the training/validation phases of the machine learning techniques. In particular, 8 folds (400 patients) were used for training and 1 (50 patients) for validation, *i.e.* parameters optimization.

TABLE 3 – The main characteristics of the CAMUS echocardiographic dataset collected from 500 patients

Dataset	Image Quality (in percentage)			LV _{EF} (in percentage)		
	Good	Medium	Poor	≤ 45%	≥ 55%	else
Full	35	46	19	49	19	32
<i>fold 1</i>	34	48	18	48	20	32
<i>fold 2</i>	34	46	20	50	18	32
<i>fold 3</i>	34	46	20	48	20	32
<i>fold 4</i>	34	46	20	50	20	30
<i>fold 5</i>	34	46	20	48	20	32
<i>fold 6</i>	36	46	18	50	20	30
<i>fold 7</i>	36	46	18	50	20	30
<i>fold 8</i>	36	46	18	50	18	32
<i>fold 9</i>	36	46	18	48	20	32
<i>fold 10</i>	36	46	18	50	18	32

3.1.2 Acquisition protocol

The full dataset was acquired from GE Vivid E95 ultrasound scanners (GE Vingmed Ultrasound, Horten Norway), with a GE M5S probe (GE Healthcare, US). No additional protocol than the one used in clinical routine was put in place. For each patient, 2D apical four-chamber and two-chamber view sequences were exported from EchoPAC analysis software (GE Vingmed Ultrasound, Horten, Norway). These standard cardiac views were chosen for this study to enable the estimation of LV_{EF} values based on the Simpson’s biplane method of discs [86]. Each exported sequence corresponds to a set of B-mode images expressed in polar coordinates. The same interpolation procedure was used to express all sequences in Cartesian coordinates with a unique grid resolution, *i.e.* $\lambda/2 = 0.3$ mm along the x-axis (axis parallel to the probe) and $\lambda/4 = 0.15$ mm along the z-axis (axis perpendicular to the probe), where λ corresponds to the wavelength of the ultrasound probe. At least one full cardiac cycle was acquired for each patient in each view, allowing manual annotation of cardiac structures at ED and ES.

3.1.3 Reference segmentation and contouring protocol

Establishing a well-defined ground-truth segmentation was of utmost importance for this work. The main difficulty when delineating 2D echocardiographic images comes from poor contrast in some regions along with the presence of well-known artifacts (*e.g.* reverberation, clutter, acoustic shadowing). One direct consequence is that embedded fully-automatic ultrasound cardiac segmentation softwares do not perform well. During the clinical exam, the clinicians delineate the different contours using semi-automatic tools under time constraints. In this context, the use of manual annotations extracted from clinical exams is not optimal to design a reference dataset for machine learning where the coherence and accuracy in the manual contouring play an important role during the learning phase.

3.1.4 Cardiologists involvement

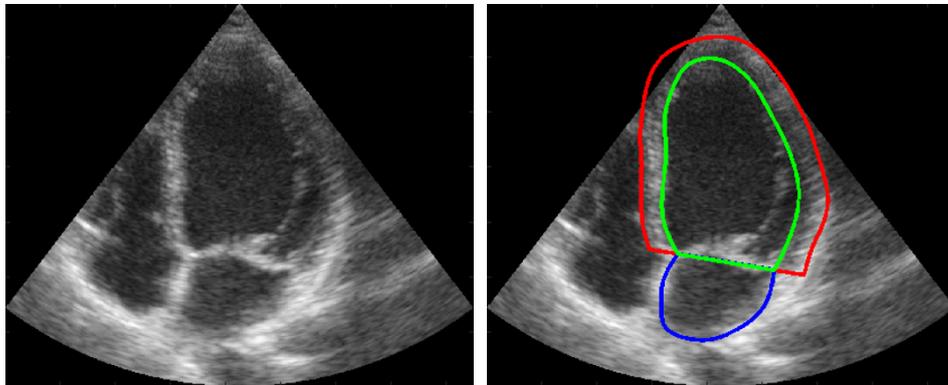
Three cardiologists (referred as O_1 , O_2 and O_3 in the sequel) participated in the annotation of the dataset. Considerable effort was spent to define a consistent manual segmentation protocol. This protocol was designed with the help of O_1 and was then strictly followed by all the involved cardiologists for the off-line manual contouring. In particular, we asked O_1 to perform the manual annotation and to determine ED and ES of the full dataset, while the two others contoured the test set of fold 5 (50 patients). O_1 also annotated twice fold 5 seven months apart (we call those annotations O_{1a} and O_{1b}). This fold was therefore used to measure both the inter- and intra-observer variability.

3.1.5 Contouring protocol

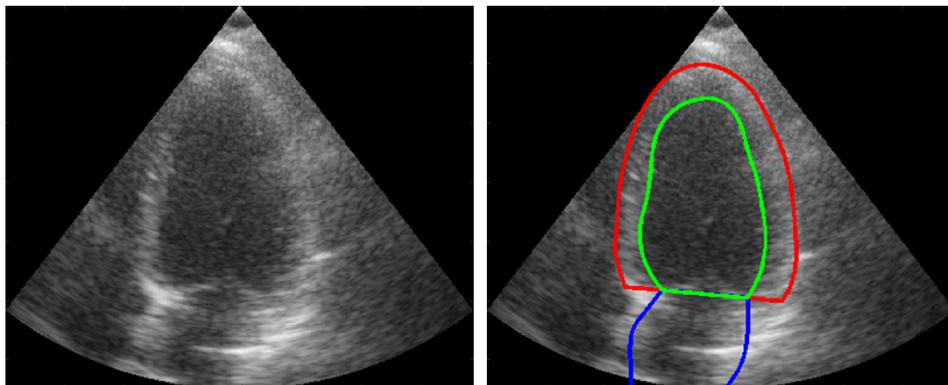
According to the recommendation of the American Society of Echocardiography and the European Association of Cardiovascular Imaging [135], ED is preferably defined as the first frame after mitral valve closure or the frame in the cardiac cycle in which the respective LV dimension or volume measurement is the largest. ES is best defined as the frame after aortic valve closure (*e.g.* using an apical long axis view image) or the frame in which the cardiac dimension or volume is smallest. In this work, ED and ES was selected as the frames where the LV dimension was at its largest or smallest, which is not the most accurate way, especially in the presence of abnormalities. This simpler approach was used due to the lack of reliable ECG. Thus the clinical indices, ED/ES volume and EF, reported in this work have to be interpreted with this in mind. While only the extraction of the LV_{Endo} contour is necessary to estimate LV_{EF} values, we also asked the cardiologists to manually outline the LV_{Epi} and the LA for all patients. This was done to study the influence of contextualization (segmentation of several structures at once) on the performance of the LV_{Endo} segmentation using deep learning techniques. The following protocol was set up.

- LV_{Endo} : Convention was used for the LV wall, mitral valve plane, trabeculations, papillary muscles and apex [135]. Basic points were to *i)* include trabeculae and papillary muscles in the LV cavity; *ii)* keep tissue consistency between ED and ES instants; *iii)* terminate the contours in the mitral valve plane on the ventricular side of the bright ridge, at the points where the valve leaflets are hinging; *iv)* partially exclude left ventricular outflow tract from the cavity by drawing from septal mitral valve hinge point to the septal wall to create a smooth shape.
- LV_{Epi} : There is no recommendation for delineating the epicardium. We thus outlined the epicardium as the interface between the pericardium and the myocardium for the anterior, anterolateral and inferior segments and the frontier between the right ventricle cavity and the septum for the inferoseptal segments.
- LA : There are recommendations for LA segmentation to assess the full LA area from dedicated LA recordings. However, since we have used acquisitions focusing on the LV, part of the dataset does not cover the full LA surface and is thus not suited to perform such measurement. Having this in mind, we used the following contouring protocol : *i)* start the LA contour from the extremities of the LV_{Endo} contour, at the points where the valve leaflets are hinging; *ii)* have the contour pass by the LA inner border.

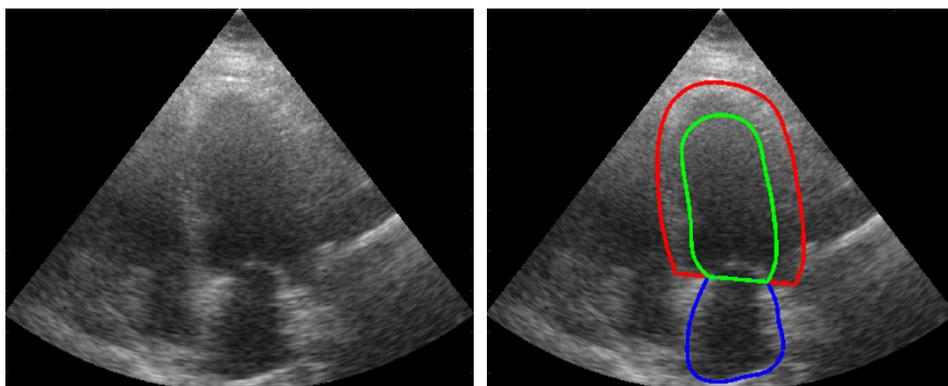
Fig. 1 illustrates our manual contouring protocol for a good, a medium, and a poor-quality image.



(a) Good image quality



(b) Medium image quality



(c) Poor image quality

FIGURE 1 – Typical images extracted from the proposed CAMUS database. Endocardium and epicardium of the left ventricle and left atrium wall are shown respectively in green, red and blue. [Left] input images ; [Right] corresponding manual annotations.

3.2 CETUS dataset

This dataset has been set-up to evaluate how far state-of-the-art methods can go at assessing 3D echocardiography, *i.e.* segmenting the LV structure as well as estimating the corresponding clinical indices. The full dataset is made available for download at <https://www.creatis.insa-lyon.fr/EvaluationPlatform/CETUS/>.

3.2.1 Patient selection

From November 2013 to August 2014, 3D images from 45 patients referred to three different hospitals (Rennes University Hospital - France, University Hospitals Leuven - Belgium and Thoraxcenter - Erasmus MC - Rotterdam - Netherlands) for 3D echocardiography were selected and included in this study (within the regulations set by the local ethical committee of each hospital). In order to provide a balanced and representative dataset of images with typical segmentation challenges, data was acquired in a population divided into 3 subgroups : 15 healthy subjects, 15 patients with a previous myocardial infarction at least 3 months prior to time of acquisition and 15 patients with dilated cardiomyopathy.

3.2.2 Acquisition protocol

In order to avoid biasing the segmentation results toward the equipment of one vendor, RT3DE exams were performed using machines from three different vendors : a GE Vivid E9, using a 4V probe, a Philips iE33, using either a X3-1 or a X5-1 probe, and a Siemens SC2000, using a 4Z1c probe. Moreover, all three hospitals acquired with two different ultrasound systems and were asked to acquire five patients from each patient group, so that patient group, hospital and ultrasound systems were equally distributed. Only images of acceptable quality for clinical diagnosis were included. The following guidelines were followed during the acquisition and inclusion of the data. Image quality had to be sufficient for visual analysis by an expert. Optimization, *e.g.* choice for harmonics, spatial resolution or other system settings were up to the operator. Exclusion criteria were *i)* left bundle branch block (LBBB) or visually dyssynchronous LV ; *ii)* frame rate below 16 volumes per second ; *iii)* mitral plane out of image sector at ED ; *iv)* significant stitching or other type of artifacts affecting the visibility of endocardium throughout the cardiac cycle ; *v)* poor visualization of a LV wall or a wall out of the image sector to an extent that the image can no longer be manually analyzed with good confidence (since one important goal of this study is the assessment of clinical indices based on volume measurements). Given that datasets were acquired in a clinical setting, a considerable variability in image quality could be anticipated (Fig. 2).

3.2.3 Training and testing datasets

Fifteen of the 45 RT3DE recordings, together with reference meshes, are made available for training or tuning of the algorithms. The remaining 30 datasets are used for testing. Care was taken to ensure that the training and testing datasets had a similar distribution of pathologies, hospitals and ultrasound machines. Acquired data were fully anonymized and handled within the regulations set by the local ethical committees of each hospital. All data was converted to a general 4D image representation format (mhd/raw) without loss of resolution. ED and ES frames were identified based on ECG and valve opening/closure by a single expert. The training dataset is released with the associated reference meshes (saved in vtk format) obtained with the contouring protocol described in Section 3.2.4.

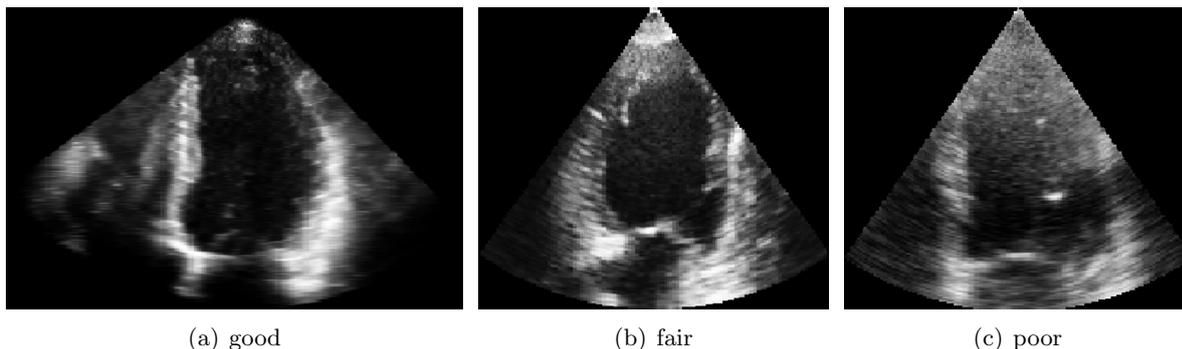


FIGURE 2 – Variability in the quality of the volumes acquired in clinical conditions

Characteristics of the training and testing datasets are given in Table 4. Image quality was assessed by one clinical expert as good, fair or poor. Image quality was slightly better in the training set. It can be seen that ED volumes as well as ES volumes are significantly different ($p < 0.01$) between the training and testing sets : the training set generally has higher volumes. These are unwanted effects of the fairly low amount of patients per subgroup. We refer the reader to [32] for more details on the acquisition setup.

3.2.4 Reference segmentation

Establishing a well-defined ground truth segmentation was of utmost importance for this work. However, there are no clear guidelines for endocardial contouring in 3D echocardiography. Therefore, considerable effort was spent to define a consistent contouring method for manual segmentation of the 3D echocardiographic data. For the ground truth drawings, we aimed for a contour definition in line with clinical standards used in 2D echocardiography. A detailed contouring guideline was set up at the beginning of the study. This guideline was refined during the training phase (contouring of the first 15 patients) and also used to resolve conflicts during consensus discussions.

Contouring protocol : We refer the reader to [190] for a complete description of the contouring procedure and protocol. In short, convention was defined for LV wall, mitral valve (MV) plane, trabeculations, papillary muscles and apex. Basic points were to a) include trabeculae and papillary muscles in the LV cavity ; b) keep tissue consistency between ED and ES instants ;

TABLE 4 – Characteristics of the two datasets which composed the CETUS database. Results as mean \pm standard deviation.* : average significantly different from the training set (unpaired t-test, $p < 0.01$).

Dataset	EDV	ESV	EF	Image Quality
	ml	ml	%	good/fair/poor
Training	213 \pm 97	151 \pm 91	33 \pm 15	6 / 6 / 3
Testing	152 \pm 62*	93 \pm 54*	41 \pm 11	8 / 10 / 12
Total	172 \pm 80	113 \pm 73	39 \pm 13	14 / 16 / 15

c) terminate the contours in the MV plane on the ventricular side of the bright ridge, at the points where the valve leaflets are hinging; d) partially exclude left ventricular outflow tract (LVOT) from the cavity by drawing from septal MV hinge point to the septal wall to create a smooth shape; e) draw the apex high up near the epicardium both in ED and ES taking into consideration that there should be little motion of the true apex point.

Tracing protocol : Manual contouring of the endocardium at the ED and ES phases was performed independently by 3 expert cardiologists from 3 distinct institutions using a custom non-commercial contouring package for 3D echocardiograms (named Speqle3D), developed by the University of Leuven and tested in an earlier study [105]. In order to minimize the impact of the reference mesh design on the segmentation error measurement, the cardiologists were asked to manually select a dense number of 3D points that belong to the endocardial border. A protocol was defined that allowed extracting this dense point cloud in a reproducible manner. At first, all datasets were oriented by defining the LV long axis, LV apex, LV base and the right ventricle insertion point. Then, each expert independently traced the endocardial border in a set of predefined planes (short axis and longitudinal planes). For each longitudinal plane about 15 points were set at the endocardial border, starting from the mitral valve plane. In the short axis planes about 10 points were set using a similar process. Finally expert meshes were generated through spherical harmonic least square fitting of the 3D point cloud. By doing so, the 3D endocardial surfaces were defined by each expert from more than 110 3D points manually annotated, which we considered to be a good trade-off between the degree of smoothness of the final endocardial shape, the accuracy of the extracted border (determined by the resolution and quality of the ultrasound equipment) and the time required for manual contouring. Since for the clinical study the volumes are computed directly from the reference meshes and since all the evaluated methods perform segmentation directly in 3D, we also decided to use the interpolated reference meshes for the computation of the segmentation accuracy.

Evaluation of correspondence and consensus : Consensus between experts was checked from pairwise differences in LV volume and EF and Hausdorff distances. To qualify for consensus between all operators the following criteria had to be met : relative difference in LV volume $\leq 10\%$, absolute difference in LVEF $\leq 5\%$ and Hausdorff distances $\leq 5mm$. All experts were asked to review the contouring of sets that did not meet criteria in the first round and suggest modifications, and one or more experts would retrace. Differences were then checked against (slightly relaxed) criteria : the average of the three pairwise observer differences was evaluated, and Hausdorff distances $\leq 7mm$ were accepted (the same initial criteria relative to the volumes and the EF having been used). In only two cases, the three operators did not agree within the consensus criteria. These drawings were then accepted, in the context of persistent observer interpretation difference. From the final contours, a mean mesh was constructed. This mean mesh (named reference mesh in the sequel) was used in the online evaluation platform as the experts' ground truth.

3.3 ACDC dataset

This dataset has been set up to evaluate how far state-of-the-art deep convolutional network methods can go at assessing cardiac magnetic resonance images (multi-slice 2D cine MRI), *i.e.* segmenting cardiac structures as well as estimating clinical indices. The full dataset is made available for download at <https://acdc.creatis.insa-lyon.fr/>.

3.3.1 Patient selection

The ACDC dataset was created from real clinical exams acquired at the University Hospital of Dijon (France). Our dataset covers several well-defined pathologies with enough cases to properly train machine learning methods and clearly assess the variability of the main physiological parameters obtained from cine-MRI (in particular diastolic volume and ejection fraction). The targeted population is composed of 150 patients evenly divided into 5 classes with well-defined characteristics according to physiological parameters. These examinations were initially classified according to medical reports. Patients with ambiguous clinical indices were excluded from this study. The different subgroups are given hereunder :

- NOR : Examination with normal cardiac anatomy and function. The ejection fraction is greater than 50%, the wall thickness in diastole is lower than 12 *mm*, the LV diastolic volume is below 90 mL/m^2 for men and 80 mL/m^2 for women [153]. The RV is normal for each patient (RV volume less than 100 mL/m^2 and RV ejection fraction above 40%). The visual analysis of the segmental LV and RV myocardial contraction is normal.
- MINF : Patients with a systolic heart failure with infarction. Subjects have an ejection fraction below 40% and abnormal myocardial contractions. Some subjects have a high diastolic LV volume due to a remodeling of the LV to compensate for the myocardial infarction.
- DCM : Patients with dilated cardiomyopathy have an ejection fraction below 40%, a LV volume greater than 100 mL/m^2 and a wall thickness in diastole smaller than 12 *mm*. As a consequence of dilated LV, some patients of this category have a dilated RV and/or a high LV mass.
- HCM : Patients with hypertrophic cardiomyopathy, *i.e.* a normal cardiac function (ejection fraction greater than 55%) but with myocardial segments thicker than 15 *mm* in diastole. In this category, patients can present abnormal cardiac mass indices with values above 110 g/m^2 .
- ARV : Patients with abnormal right ventricle have a RV volume greater than 110 mL/m^2 for men, and greater than 100 mL/m^2 for women [160], or/and a RV ejection fraction below 40%. Almost every subject in this sub-group has a normal LV.

3.3.2 Acquisition protocol

Acquisitions were obtained over a 6 year period with two MRI scanners of different magnetic strengths (1.5 T - Siemens Area, Siemens Medical Solutions, Germany and 3.0 T - Siemens Trio Tim, Siemens Medical Solutions, Germany). Cine MR images were acquired with a conventional SSFP sequence in breath hold with a retrospective or prospective gating [235]. After the acquisitions of long axis slices, a series of short-axis slices covering the LV from the base to the apex was acquired, with a slice thickness from 5 *mm* to 10 *mm* (in general 5 *mm*) and sometimes an inter-slice gap of 5 *mm*. The spatial resolution varies from 1.34 to 1.68 $mm^2/pixel$. Depending on the patient, 28 to 40 volumes were acquired to cover completely (retrospective gating) or partially (prospective gating) one cardiac cycle. In the latter case, only 5 to 10% of the end of the cardiac

cycle was omitted. The full dataset was acquired in clinical routine, leading to natural variability in the image quality (intrinsic noise, patient movement, banding artifacts, MRI low-frequency intensity fluctuation, etc.), variable field-of-view and integral or almost integral covering of the LV. Finally, to be in compliance with previous cardiac MRI segmentation challenges, the long axis slices were not provided. Even though the use of long axis slices could provide extra information about the base, the apex and the longitudinal motion of the ventricles, the analysis of short and long-axis slices are generally independent and outside the scope of this project.

3.3.3 Training and testing datasets

The data for each subject was converted to a general 4D image representation format (nifti) without loss of resolution. ED and ES frames were identified based on the motion of the mitral valve from the long axis orientation by a single expert. Both training and testing data contain whole short-axis slices. The identification of the most basal and apical slices is also not provided, while the diastolic and systolic phases are indicated. In order for challengers to normalize the physiological parameters (mainly the LV and RV volumes and the MYO mass) with the body surface area (BSA), the weight and height of each patient are included in the dataset. For instance, the BSA can be calculated from the formula of Dubois and Dubois [74], *i.e.* $BSA = 0.007184 \cdot (weight^{0.425} \cdot height^{0.725})$ and normalized parameters can be computed by simply dividing their values with the corresponding BSA. The training database is composed of 100 patients, *i.e.* 20 patients for each group. For all these data, the corresponding manual references as well as the patient group are provided. The testing dataset is composed of 50 patients, *i.e.* 10 patients per group. The manual references and group labels of the testing data are kept private.

3.4 Reference segmentation and contouring protocol

The expert references are manually-drawn 3D volumes of the LV and RV cavities as well as the myocardium, both at the ED and ES gates. The epicardial border of the RV was not considered because its accurate position next to the septum is difficult to establish, and the myocardial thickness of the RV is of the same order of magnitude than the spatial resolution. The contours were drawn and double-checked by two independent experts (10 and 20 years of experience) who had to reach consensus in case of discordance. The following annotation rules were retained : the LV and RV must be completely covered, the papillary muscle are included into the cavity and there is no interpolation of the muscle at the base of the LV (the contours follow the limit defined by the aortic valve). The main difficulty when annotating RV corresponds to correctly localize the pulmonary infundibulum area. This area must not be included into the RV annotation and a clear separation must be seen between the RV cavity and the root of the pulmonary artery. Due to the systolic shortening of the RV, the first basal slice is not mandatory being the same in diastole and systole. Another difficulty is to accurately separate the RV from the right atrium on the systolic image. As such, we defined the RV as the region on the right of heart with a significant contraction between ventricular diastole and systole, *i.e.* the surface area of the RV must be higher in ventricular diastole than in ventricular systole. For an easier understanding, illustrations of the annotation rules are provided in the supplementary materials (available in the supplementary files /multimedia tab). The ground truth label images were stored in nifti format. The label values vary from 0 to 3 and represent voxels belonging to the background (0), the RV cavity (1), the myocardium (2) and the LV cavity (3). Fig. 3 to 4 illustrate the protocol used for the manual delineation to create the segmentation references.

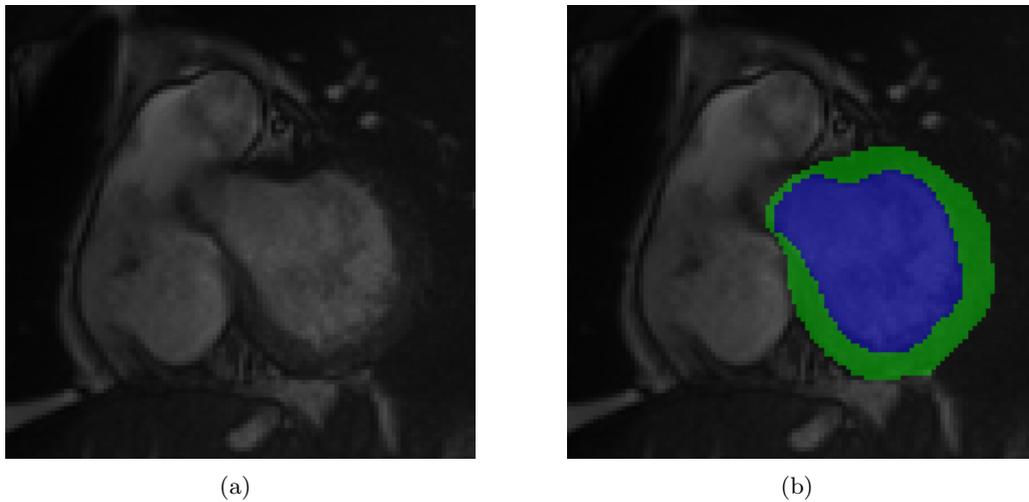


FIGURE 3 – Manual segmentation of the left ventricle at the level of the most basal slice. a) Original image. b) Image with the left ventricle (in blue) and the myocardium (in green).



FIGURE 4 – Manual segmentation of the right ventricle at the level of the most basal slice. a) Original image. b) Right ventricle segmentation in red. Only the posterior part was considered because of the absence of contraction at the level of the anterior area (white arrow) between diastole and systole (it is thus not considered as a part of the ventricle).

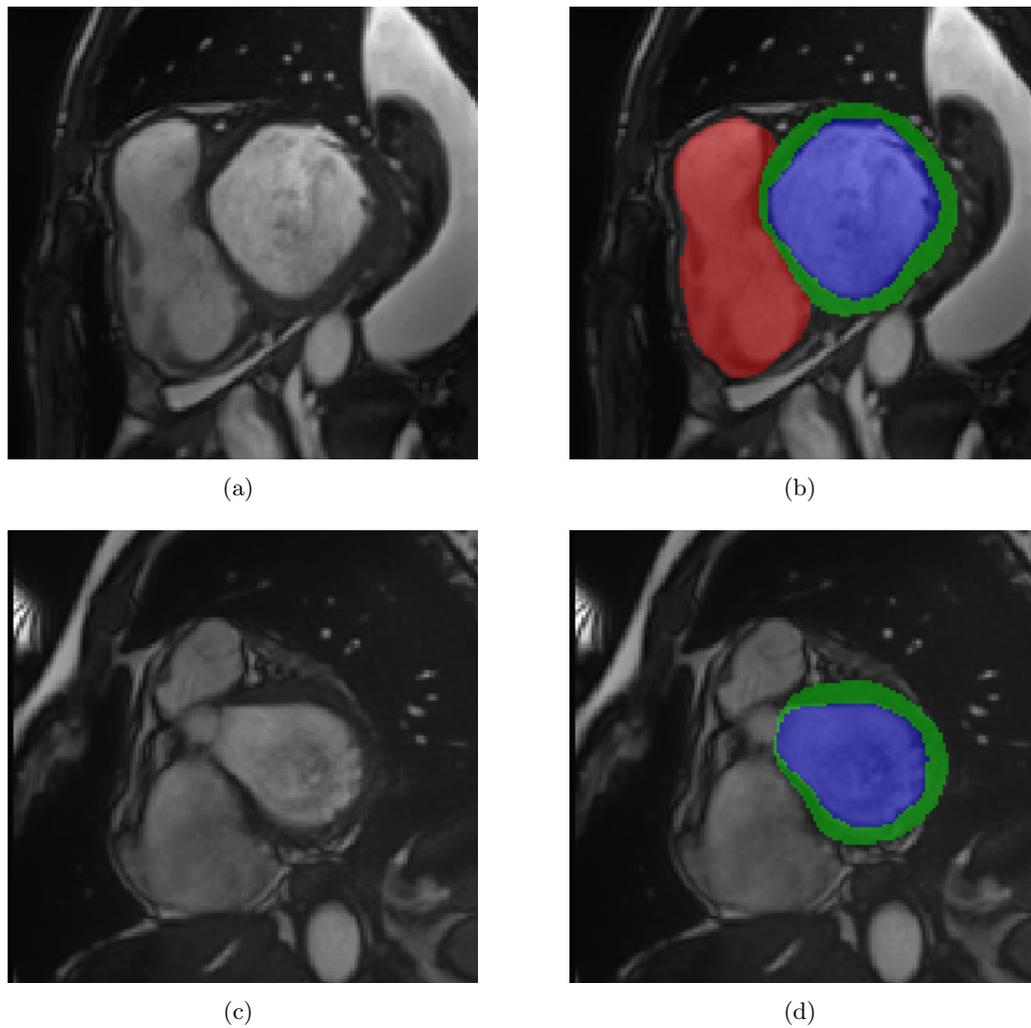


FIGURE 5 – Manual segmentation of the right ventricle (RV) at the level of the most basal slice. The left ventricle is in blue, the myocardium in green and the right ventricle in red. The first column displays original images while each row corresponds to a given patient. b) RV's border is drawn when there is no clear separation between the cavity and the aorta. d) No RV's border is drawn when there exists a clear separation between the cavity and the aorta.

4 2nd contribution : B-spline Explicit Active Surface paradigm (BEAS)

In this section we present a novel framework, named B-spline Explicit Active Surface, that we introduced to efficiently segment RT3DE. It is based on an efficient explicit representation of a deformable model and a dedicated automatic initialization scheme. The obtained results were evaluated through the CETUS database described in Sec. 3.2. This work has been the subject of several publications [9, 22, 23, 25, 30, 197].

4.1 Methodology

4.1.1 BEAS framework

The key concept of the BEAS framework [25] is to regard the boundary of an object as an explicit function, where one of the coordinates of the points on the surface, $\mathbf{x} = \{x_1, \dots, x_n\}$, is given explicitly as a function of the remaining coordinates, *i.e.* $x_1 = \psi(x_2, \dots, x_n)$. In this framework, ψ is defined as a linear combination of B-spline basis functions :

$$x_1 = \psi(x_2, \dots, x_n) = \psi(\mathbf{x}^*) = \sum_{\mathbf{k} \in \mathbb{Z}^{n-1}} c[\mathbf{k}] \beta^d \left(\frac{\mathbf{x}^*}{h} - \mathbf{k} \right), \quad (1)$$

where \mathbf{x}^* is the point of coordinates $\{x_2, \dots, x_n\}$ and $\beta^d(\cdot)$ the uniform $(n-1)$ -dimensional B-spline of degree d . The knots of the B-splines are located on a rectangular grid defined on the chosen coordinate system, with a regular spacing given by h . The coefficients of the B-spline representation are gathered in $c[\mathbf{k}]$.

Given the volumetric nature of the left ventricle, the B-spline representation was created on a spherical coordinate system thus defining the active geometric functions as $r = \psi(\phi, \theta)$. The angular discretization of the boundary representation was set empirically at 24×16 (elevation \times azimuth) and the B-spline scale to 2^1 for both angular coordinates.

The evolution of the model is defined by the minimization of an energy criterion E . This energy is expressed by the sum of the data attachment term E_d and a regularization term E_r :

$$E = E_d + E_r. \quad (2)$$

The data attachment energy function E_d follows a variation of the localized Yezzi energy adapted for endocardial segmentation [22], thus taking into account the expected intensities of the blood pool and the endocardium :

$$E_d = \int_{\Omega} \delta_{\phi}(\mathbf{x}) \int_{\Omega} B(\mathbf{x}, \mathbf{y}) \cdot (u_{in} - u_{out}) dy d\mathbf{x}, \quad (3)$$

where $\delta_{\phi}(\mathbf{x})$ is the Dirac operator applied to the level set function $\phi(\mathbf{x}) = \psi(\mathbf{x}^*) - x_1$, which is defined over the image domain Ω and where $\Gamma = \mathbf{x} \in \mathbb{R}^n \mid \psi(\mathbf{x}^*) = x_1$ represents the surface being segmented. u_{in} and u_{out} are the local intensity means around x , respectively inside and outside the surface. $B(\mathbf{x}, \mathbf{y})$ is the mask function in which these local parameters are estimated, restricted to the points along $\mathbf{N}(\mathbf{x})$, the normal direction of the surface, at a distance smaller than ρ :

$$B(\mathbf{x}, \mathbf{y}) = \begin{cases} 1, & \text{if } \mathbf{y} = \mathbf{x} + k\mathbf{N}(\mathbf{x}), k \in [-\rho, \rho] \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

The neighborhood region limit ρ was set experimentally at 16 mm. The minimization of the data attachment energy term in Eq. (3) can then be performed through optimization of the B-spline coefficient $c[\mathbf{k}_i]$, thus :

$$\frac{\partial E_d}{\partial c[\mathbf{k}_i]} = \int_{\Gamma} \left(\frac{\bar{I}(\mathbf{x}^*) - u_{in}}{A_{in}} + \frac{\bar{I}(\mathbf{x}^*) - u_{out}}{A_{out}} \right) \beta^d \left(\frac{\mathbf{x}^*}{h} - \mathbf{k}_i \right) d\mathbf{x}^*, \quad (5)$$

where A_j is the area of region j used to estimate the local mean u_j and $\bar{I}(\mathbf{x}^*)$ corresponds to the image value at the position $\mathbf{x} = \{\psi(\mathbf{x}^*), x_2, \dots, x_n\}$.

4.1.2 Statistical shape model

In order to provide accurate shape information to the proposed BEAS approach, a sufficiently broad and numerous dataset of 3D LV shapes is needed. For that purpose, 289 cardiac magnetic resonance (cMR) datasets from a large multi-center clinical study, DOPPLER-CIP [216], were used. This study was aimed at patients whose profile corresponds to suspected chronic ischemic disease and thus encompasses patients of a broad clinical spectrum. The cMR datasets were contoured by experts at ED and ES and the 2D slices were aligned using an iterative closest point (ICP) algorithm [37] to correct for breath-hold slice misalignment. A 3D mesh was then interpolated from the aligned 2D contours at ED and ES for each patient.

Similar to Queirós *et al.* [212], the statistical shape model (SSM) was built in the BEAS coordinate system ; in this case in spherical coordinates. The SSM shapes will then be represented by their B-spline representation coefficients $c[\mathbf{k}]$. Because such a representation assumes that the position and orientation of the coordinate system is identical for every shape, the position and orientation of the training shapes have to be aligned, which can be done according to the centroid and direction of largest variance of each shape.

Starting from the aligned 3D LV shapes in BEAS space, the first step to build the SSM is to scale all shapes so that equivalent points from different shapes can be compared without the influence of the LV size. Considering $c_s[\mathbf{k}]$ the s^{th} shape of all N shapes, this is done by : calculating the mean of all shapes $\bar{c}[\mathbf{k}] = \frac{1}{N} \sum_{s=1}^N c_s[\mathbf{k}]$, scaling each shape to the current mean $\bar{c}[\mathbf{k}]$ and then repeating these steps until the process converges [51]. The scaling step is done according to :

$$c_{scaled}[\mathbf{k}] = c[\mathbf{k}] \frac{\sum_i w[\mathbf{k}_i] \bar{c}[\mathbf{k}_i] c[\mathbf{k}_i]}{\sum_i w[\mathbf{k}_i] c[\mathbf{k}_i] c[\mathbf{k}_i]}, \quad (6)$$

where $w[\mathbf{k}]$ is a set of weights chosen to give more significance to the points that tend to be most stable :

$$w[\mathbf{k}_i] = \left(\sum_{s=1}^N \text{Variance}(c_s[\mathbf{k}_i]) \right)^{-1}. \quad (7)$$

Principal Component Analysis (PCA) can then be applied to extract the shape variability of the LV B-spline coefficients [51]. Through singular value decomposition [122], it is then possible to obtain the eigenvectors \mathbf{p}_i and the corresponding eigenvalues λ_i of the covariance matrix :

$$\mathbf{S} = \frac{1}{N} \sum_{s=1}^N (c_s[\mathbf{k}_i] - \bar{c}[\mathbf{k}_i]) (c_s[\mathbf{k}_i] - \bar{c}[\mathbf{k}_i])^T. \quad (8)$$

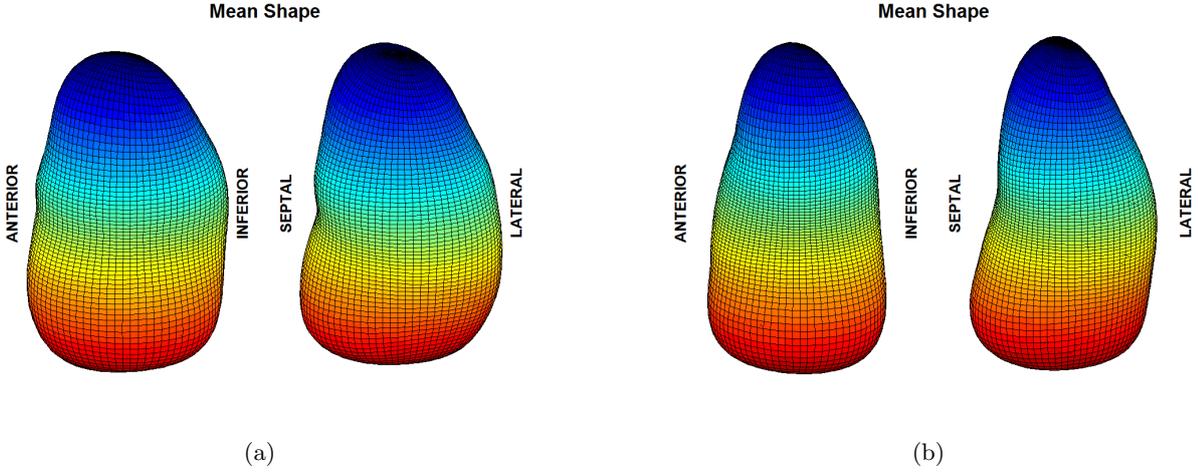


FIGURE 6 – Mean statistical shape models at ED (a) and ES (b).

Since most of the variation can be explained by a small number of eigenvectors, only a portion of the original set is kept, corresponding to the number of eigenvectors t whose sum represents 90% of the total variance of all variables. In this way, any shape from the dataset can be approximated by :

$$c[\mathbf{k}] \approx \bar{c}[\mathbf{k}] + \mathbf{P}\mathbf{b}, \quad (9)$$

where \mathbf{P} is the matrix of the first t eigenvectors and \mathbf{b} is a vector of t weights which for any given shape corresponds to $\mathbf{b} = \mathbf{P}^T (c[\mathbf{k}] - \bar{c}[\mathbf{k}])$.

To be able to model both ED and ES separately, two different SSMs were created according to the methodology described above. The mean shapes for each of these models are shown in Fig. 6. Note that since these models are scaled according to Eq. (6) only shape variations can be observed in this figure.

4.1.3 SSM-Based Regularization

To then use the SSM with BEAS for the segmentation of new images, two different regularization energies were defined so that the segmented shapes fit those observed in the training set. These two energies, a hard and a soft SSM-based regularization, were first proposed by Queirós *et al.* [212]. In the present study, those energies were adapted to regularize the B-spline coefficients $c[\mathbf{k}]$ of a 3D shape.

The hard SSM-based regularization restricts the segmented shape to the shape variability observed in the training set. At each iteration, the weights \mathbf{b} are computed and each is restricted to $\pm m\sqrt{\lambda_i}$, where m defines the limits to the variability from the mean [51]. m is typically set between 2 and 3 since most of the population lies within three standard deviations and was set at 2.5 in this study. Through Eq. (9), a new regularized shape $c_{reg}[\mathbf{k}]$ is then obtained [212]. To include this hard SSM-based term in the BEAS framework, the energy functional can be defined as :

$$E_{hard} = \int_{\Omega} \delta_{\phi}(\mathbf{x}) \left(\frac{1}{2} (c[\mathbf{k}] - c_{reg}[\mathbf{k}])^2 \right) d\mathbf{x}. \quad (10)$$

The minimization of this energy according to the B-spline coefficients gives :

$$\frac{\partial E_{hard}}{\partial c[\mathbf{k}]} = \int_{\Gamma} (c[\mathbf{k}] - c_{reg}[\mathbf{k}]) d\mathbf{x}^*. \quad (11)$$

The soft SSM-based regularization follows the rationale that it is much more probable to find an average shape than a shape which is close to the variability limits. In that way, the soft SSM-based regularization penalizes high values of b_i and is defined as the squared Mahalanobis distance to the training shapes [51], thus :

$$E_{soft} = \sum_{i=1}^t \frac{b_i^2}{\lambda_i} = \int_{\Omega} \delta_{\phi}(\mathbf{x}) \left[(c[\mathbf{k}] - \bar{c}[\mathbf{k}])^T \mathbf{S}^{-1} (c[\mathbf{k}] - \bar{c}[\mathbf{k}]) \right] d\mathbf{x} \quad (12)$$

Following the derivation shown in Queirós et al. [212], the minimization of E_{soft} gives :

$$\frac{\partial E_{soft}}{\partial c[\mathbf{k}]} = \int_{\Gamma} 2\mathbf{P}\mathbf{D}^{-1}\mathbf{b} d\mathbf{x}^*, \quad (13)$$

where \mathbf{D} is the diagonal matrix of t eigenvalues λ . To incorporate these two energies into BEAS, the regularization term E_r is defined as :

$$E_r = \alpha E_{hard} + \beta E_{soft}, \quad (14)$$

where α and β are hyperparameters controlling the relative weight between the two terms.

4.2 Automatic initialization strategy

The key steps of the proposed algorithm are the sequential detection of the left ventricular long axis (LAX) and mitral valve plane. Once this information is available, we take advantage of the information gathered in these two first steps to fit an ellipsoid to the detected endocardial boundaries. The overview of the key processing blocks of the algorithm are shown in Fig. 8 and 7.

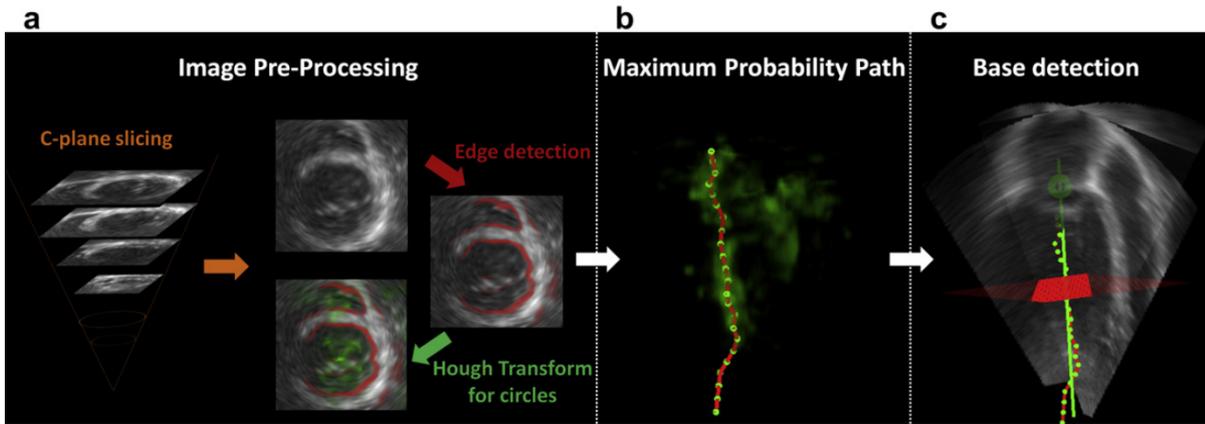


FIGURE 7 – Left ventricular long axis (LAX) and base detection : (a) Image pre-processing (red - edge indicator, green - circle center probability); (b) Maximum probability path, displayed over P_{max} ; (c) Base detection (red - sliding averaging plane; green ball detected base position).

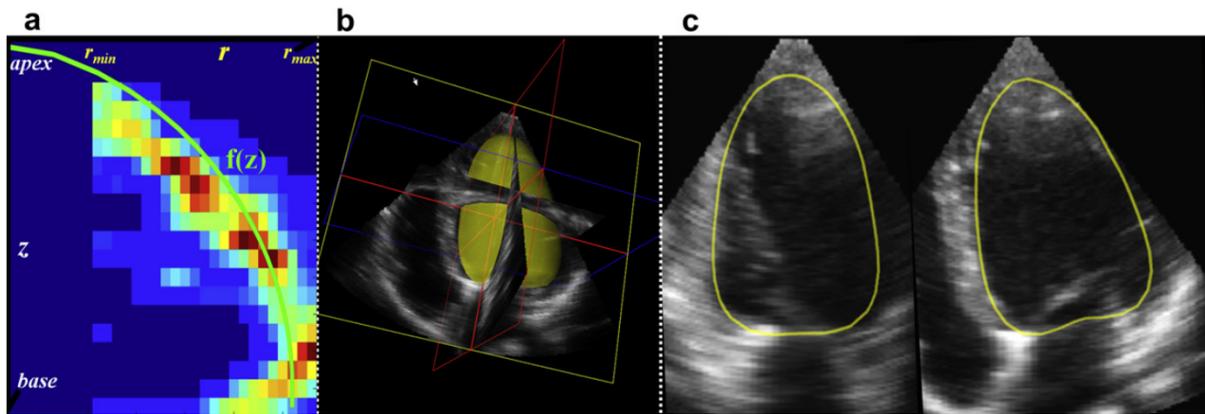


FIGURE 8 – Ellipsoid fitting : (a) Feature map and optimal ellipsoid fit in cylindrical coordinates (green); (b and c) Initial surface for the 3-D image segmentation step.

4.2.1 Image preprocessing

The first step for the automatic initialization of a 3-D ultrasound volume is to detect the myocardial boundaries using edge operators in several C-planes (*i.e.* planes parallel to the transducer surface), perpendicular to the acquisition axis. Given that ultrasound images possess low contrast between the blood pool and the myocardial tissue, it is proposed to include a robust, phase-based edge detector, because of its intrinsic invariance to intensity. The local phase extracted using monogenic signal and applied to 3-D ultrasonic data by Rajpoot *et al.* [217] has a remarkable performance in the detection of the myocardial boundaries in low quality ultrasound data, even in low contrast areas. This edge detector is applied to each of the 2-D C-planes taken 5 mm apart.

Since the endocardium is approximately a circle in these 2-D C-planes, low level computer vision methods can be employed to robustly find the most likely position for the center of the LV cavity. To this end, the Hough transform for circles (HTc) is employed. This transform will return

the probability of a given position (x, y) in the 2-D C-plane image being the center of a circle of radius r . By concatenating the output of the HTc, a 4D (x, y, z, r) circle center probability matrix, P_c , can be generated. To keep r in a physiologically relevant range, r_{min}, r_{max} was set to [15, 35] mm. A schematic diagram of the steps taken in the image preprocessing block are shown in Fig. 7-a. The multiscale phase-based edge detector algorithm used wavelengths of 24, 30 and 36 pixels, as suggested in the original submission by Rajpoot *et al.* [217]. Its result was normalized to its maximum value and thresholded using $Th = 0.05$ to obtain a binary edge map detector prior to circle detection.

4.2.2 Long axis detection

To reduce the computational complexity of the long axis detection step, P_c was transformed into a 3-D matrix, P_{max} , by taking into account only the maximum probability value along the r direction. Subsequently, multi-dimensional dynamic programming (MDP) was used to find the optimal path that maximizes the center probabilities along the z direction [270]. For each C-plane slice, P_{max} can be seen as a density map and thus, thanks to the MDP algorithm, a path corresponding to a set of positions on each of the 2-D C-planes that have high probability of being a center of a circle will be found. Note that during the MDP estimation of the optimal path, connectivity constraints are imposed to penalize the variation of (x, y) coordinates, to have a smoother path. Through the positions of the path found with MDP, a straight line is fit to determine the left ventricular long axis. A demonstrative example of P_{max} (in green transparency) and the optimal path found with MPD (green dots connected with red segments) are shown in Fig. 7-b.

4.2.3 Base detection

After LAX detection, a perpendicular square plane of size r_{max} was slid along the LAX. For each position, an average gray level was taken by averaging intensities over this square plane. Given that the base plane corresponds both to an area of brighter appearance and to a dark-to-bright transition, a simple merit function, whose maximum corresponds to the base position, was built as :

$$g(z) = I_{LAX}(z) + (I_{LAX}(z) - I_{LAX}(z - \zeta)) \quad (15)$$

where $I_{LAX}(z)$ corresponds to the average gray level of the square plane perpendicular to the LAX at depth z and ζ is a distance parameter used to assess the expected dark-to-bright transition. ζ was set to 0.5 cm in all the experiments. Note that the first term in Eq. (15) accounts for the brighter appearance of the valve plane and the second term accounts for the variation in the gray level intensity attributable to the blood-valve plane transition.

4.2.4 Ellipsoid fitting

P_c was sampled along the estimated LAX resulting in a 2-D feature map, $FM(z, r)$, which will correspond to the probability of the estimated LAX positions being the center of a circle of radius r . Next, this feature map needs to be converted to a true 3-D model that can be used to initialize the segmentation algorithm. To this end, we propose to formulate the ellipsoid fitting problem using its parametric equation defined in a cylindrical coordinate system. It should also

be noted that all the required data was already computed in the long axis detection step, being thus a very computationally efficient solution. Consider the ellipsoid parametric equation :

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1. \quad (16)$$

Assuming that $a = b$ (*i.e.* that the ellipsoid is actually a spheroid), it can be easily shown that $x^2 + y^2$ is actually the radius of the circle that results from the intersection of the ellipsoid and a plane perpendicular to the z axis. Thus, the radius of each z -cut of the ellipsoid can be expressed as :

$$r = f(z) = a\sqrt{1 - \frac{z^2}{c^2}}. \quad (17)$$

The apex was fixed at 2 cm from the probe to provide a rough estimate of the longitudinal measure for the LV-ellipsoid fitting process. From all the plausible curves $f(z)$ ($a \in [r_{in}, r_{max}]$ and $c \leq d_{LAX}$, where $d_{LAX} = base_Z - 2cm$), the one maximizing the sum of the radius probability along the LAX was selected, as shown in Fig. 8-a. Once an automatically fitted ellipsoid is available for the end-diastolic (ED) frame, it will be used as initialization of the segmentation algorithm.

4.2.5 Automatic SAx orientation

The short axis (SAx) orientation method used was proposed in Pedrosa *et al.* [196]. This method aims at the detection of the right ventricular inferior insertion point and relies on image intensity information and analysis of the structures surrounding the LV. A detailed description of this method can be found in Pedrosa *et al.* [196]. The SAx orientation is crucial to correctly position the SSM, given that different sides of the LV have different shape characteristics. However, this orientation algorithm depends on a previous estimate of the LV surface and the result from the initialization is too rough as it relies on the fitting of an ellipsoid. As such, the automatic SAx orientation is only applied after an initial stage of segmentation with BEAS.

4.3 Framework description

A conceptual description of the proposed framework is shown in Fig. 9.

4.3.1 Segmentation at ED

The segmentation at ED is composed of two stages. Initially, BEAS is used without the SSM, so that the energy criterion E is equal to E_d , the data attachment term. This provides an initial segmentation of the LV, which is used for the SAx orientation estimation but also to refine the initial estimates of LAX orientation and center according to the centroid and direction of largest variance of the segmented mesh. With the center position and both the LAX and SAx orientation well defined, it is then possible to use BEAS with the SSM regularization according to Eq. (14) to further refine the segmentation.

4.3.2 Localized Anatomical Affine Optical Flow (IAAOF)

IAAOF is then used to propagate the result from ED to the remaining frames. The IAAOF method was proposed in Queirós *et al.* [214] and relies on an affine optical flow approach which independently estimates the motion at each point in the surface based on an anatomically

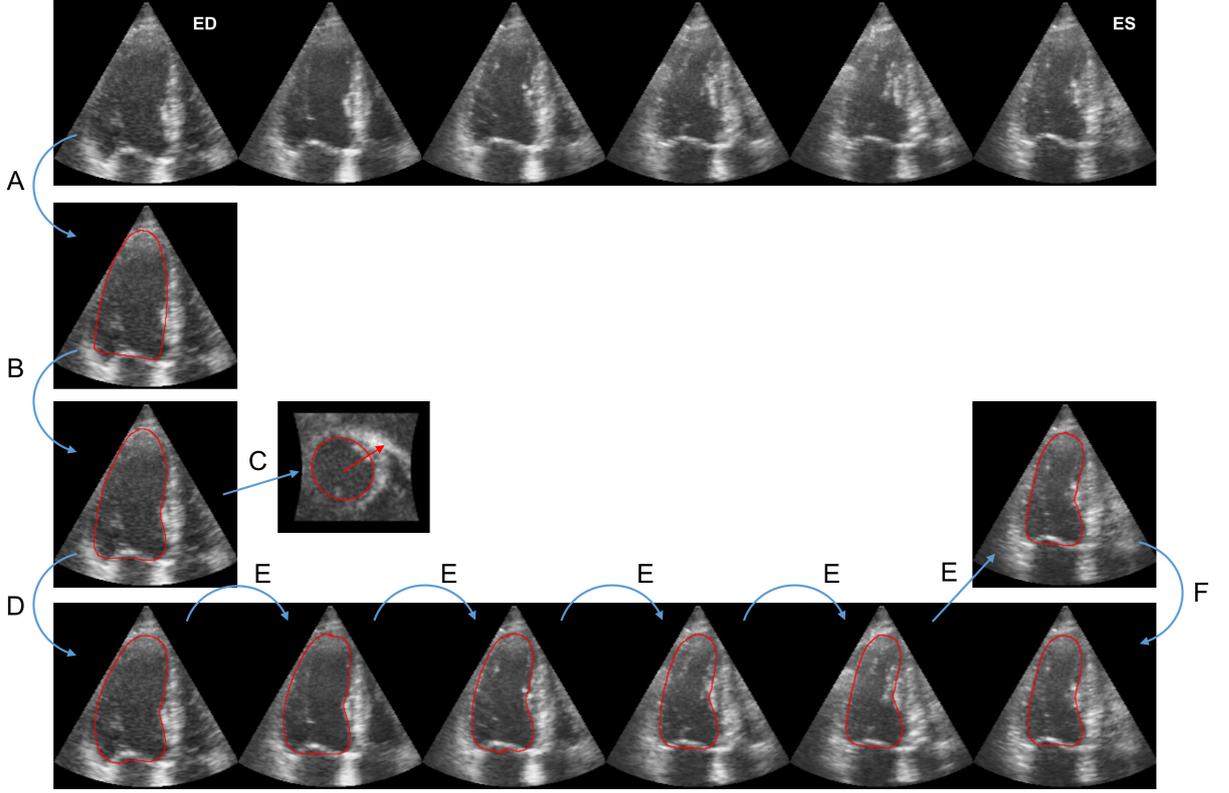


FIGURE 9 – Conceptual description of the proposed segmentation and tracking framework. First, automatic initialization is applied to the ED frame (A). The first stage of segmentation is then performed using BEAS (B). The result from this segmentation is used to detect the SAX orientation (C) and this information is then used to perform the second stage of segmentation using BEAS and the ED SSM. The final ED segmentation is then propagated frame to frame using the IAAOF (E) and a final refinement to the ES frame is performed using BEAS and the ES SSM (F).

constrained neighborhood. A detailed description of this method can be found in the original paper by Queirós *et al.* [214]. The parameters used to tune the IAAOF were replicated from [214].

4.3.3 Segmentation at ES

Segmentation at ES is used to further refine the result from the IAAOF, thus bringing together intensity and shape-based clues. In order to balance the contribution between tracking and segmentation clues, an energy term was added to penalize the deviation between the result of the IAAOF and the segmentation. Such an approach was first proposed by Barbosa *et al.* in [21] and can be formulated as :

$$E_A = \int_{\Gamma} \left(\psi(\mathbf{x}^*) - \tilde{\psi}(\mathbf{x}^*) \right)^2 d\mathbf{x}^*, \quad (18)$$

where $\tilde{\psi}(\mathbf{x}^*)$ is the surface obtained from the tracking using IAAOF. The minimization of

this energy with regard to the B-spline coefficients can be performed according to :

$$\frac{\partial E_A}{\partial c[\mathbf{k}]} = \int_{\Gamma} 2 \left(\psi(\mathbf{x}^*) - \tilde{\psi}(\mathbf{x}^*) \right) \beta^d \left(\frac{\mathbf{x}^*}{h} - \mathbf{k}_i \right) d\mathbf{x}^*. \quad (19)$$

The regularization energy criterion is then expressed as :

$$E_r = \alpha E_{hard} + \beta E_{soft} + \gamma E_A, \quad (20)$$

where γ is a hyperparameter used to define the balance between tracking and intensity/shape-based information.

4.4 Experiments

4.4.1 Data description

The proposed framework was tested on the CETUS database described in Sec. 3.2. In particular, fifteen datasets with the corresponding reference meshes at ED and ES were used as training, while the remaining 30 datasets were used as test set.

4.4.2 Segmentation performance assessment

First, the 15 training datasets were used to tune the hyperparameters β and γ needed respectively for the SSM regularization and for the balance between the segmentation and tracking information. This tuning was performed empirically by visual inspection of the results. The hyperparameters α , β and γ were set respectively to 1, 0.0005 and 0.25. Note that the value of β is directly related to the absolute value of eigenvalues λ as defined in Eq. (12), thus justifying its relative small value.

Using these settings, the framework was then tested on the 30 testing datasets. The evaluation of the results was conducted using the online MIDAS platform of the CETUS challenge, thus assuring that the proposed method can be directly compared to other state-of-the-art methods. The accuracy of the segmentation was evaluated at ED and ES through different distance metrics : Mean Absolute Distance (MAD) [44], which measures the average distance at any point between the segmented and reference meshes ; Hausdorff Distance (HD) [109], which measures the maximum distance between the segmented and reference meshes ; and Dice [64], which is a measure of the overlap between the segmented and reference meshes. Because the meshes obtained from BEAS are sampled in the spherical coordinate system, causing the point density to be different along the surface, which could bias the error metrics to specific regions, the segmented meshes were remeshed to assure greater smoothness and more uniform mesh point density. Clinical indices were also studied, namely the Pearson correlation coefficient and limits of agreement of ED volume (EDV), ES volume (ESV) and ejection fraction (EF). Mean computational times of the proposed framework were also obtained using MATLAB code running on an Intel® Xeon® E5-1650v2@3.5GHz with 32GB RAM.

4.4.3 Position/orientation performance and sensitivity

Because the characteristics of the SSM are closely related to the position and orientation (LAX and SAX) of the BEAS coordinate system, it is important to determine the error in the estimation of these parameters. For that purpose, the position, LAX orientation and SAX

orientation of the CETUS training set reference meshes were compared to those obtained with the proposed method.

Furthermore, the sensitivity of the segmentation results to variations of these parameters was studied. This was performed by introducing variations from the reference position or orientation on each of these parameters and evaluating the segmentation performance. In this way, to evaluate, for example, the sensitivity to the position, BEAS was initialized at a random position at a distance D from the reference mesh position and with the reference SAx and LAx orientation. The segmentation result was then evaluated on MAD, HD and Dice. To prevent sporadic results from this random positioning, each image was started from three different random positions each time and the results averaged.

4.4.4 Parameter sensitivity assessment

To study the robustness and stability of the proposed framework with respect to the multiple parameters involved, a parameter sensitivity assessment was conducted. As such, the balance of the different energies, namely α , β and γ , was studied. Each parameter was varied from their empirically determined preset by 50% of its value and its impact studied in terms of MAD, HD and Dice. To further analyze the contribution of each component of the framework, the segmentation performance was analyzed when each of these energy parameters was set to zero. To highlight the importance of the IAAOF, the segmentation performance of the framework without the IAAOF was also studied by using the ED segmentation result for initialization of the ES segmentation.

4.4.5 Statistical analysis

Paired t-tests were used to analyze the significance of differences between the proposed method and other methods in literature and to analyze the parameter sensitivity of the proposed method. Results are denoted as mean \pm standard deviation.

4.5 Results

4.5.1 Segmentation performance

Tables 5 and 6 show the segmentation and tracking results for the proposed approach, as well as the performance obtained with other state-of-the-art methods and inter-observer variability from manual contouring. Those obtained by Queirós *et al.* [214] and Barbosa *et al.* [27] also use BEAS as the segmentation tool but neither use shape-based information. Queirós *et al.* used the same IAAOF tracking whereas Barbosa *et al.* used a global anatomically constrained optical flow approach followed by block matching refinement instead of the IAAOF. The other approaches presented were chosen as they are, to the author's knowledge, the ones presenting the best segmentation results on the CETUS dataset.

A regionwise analysis of error was also conducted by dividing the LV into the 17-segment model [42], using the LAx as reference and dividing the LV into basal (35%), mid-cavity (35%) and apical (30%) regions. The average MAD and HD at ED and ES for the training datasets is shown in Fig. 10. It can be observed that the greatest errors occur on the apical region and on the anterior side of the LV. Fig. 11 shows examples of the fully automatic segmentation results compared to the consensus manual contours by experts.

TABLE 5 – Performance on the CETUS testing datasets. MAD, HD and Dice of the proposed framework, other state-of-the-art approaches and inter-observer variability. All values in mean \pm standard deviation (*NR* stands for not reported). \star , \dagger and \ddagger indicate respectively that the difference to the proposed framework was statistically significant at a $p < 0.05$, $p < 0.01$ and $p < 0.001$ level. Note that for methods [36, 184, 190, 272] a comparison is not possible as the data is not publicly available.

Methods	MAD (mm)		HD (mm)		Dice	
	ED	ES	ED	ES	ED	ES
BEAS	1.81 ± 0.59	1.98 ± 0.66	6.31 ± 1.69	6.95 ± 2.14	0.909 ± 0.034	0.875 ± 0.046
Queirós <i>et al.</i> [214]	2.26 \ddagger ± 0.72	2.45 \ddagger ± 0.85	8.10 \ddagger ± 2.62	8.19 \star ± 3.03	0.894 \ddagger ± 0.040	0.861 \star ± 0.054
Barbosa <i>et al.</i> [27]	2.26 \ddagger ± 0.72	2.43 \ddagger ± 0.89	8.10 \ddagger ± 2.62	8.29 \star ± 3.01	0.894 \ddagger ± 0.040	0.856 \ddagger ± 0.056
Bernier <i>et al.</i> [36]	2.37 <i>NR</i>	2.64 <i>NR</i>	9.41 <i>NR</i>	9.34 <i>NR</i>	0.882 <i>NR</i>	0.837 <i>NR</i>
van Stralen <i>et al.</i> [272]	1.91 <i>NR</i>	2.48 <i>NR</i>	6.66 <i>NR</i>	7.38 <i>NR</i>	0.910 <i>NR</i>	0.862 <i>NR</i>
Oktay <i>et al.</i> [184]	1.94 ± 0.55	2.23 ± 0.60	7.00 ± 1.99	7.53 ± 2.23	0.904 ± 0.02	0.874 ± 0.04
Inter-observer variability <i>et al.</i> [190]	1.01 ± 0.30	1.01 ± 0.38	3.37 ± 0.87	3.30 ± 0.94	0.949 ± 0.15	0.938 ± 0.21

TABLE 6 – Performance on the CETUS testing datasets. Pearson correlation coefficient (R) and limits of agreement (LOA) (mean \pm standard deviation) in comparison to the reference cardiac indices (EDV, ESV and EF) of the proposed framework, other state-of-the-art approaches and inter-observer variability. Volumes obtained with the proposed method were not statistically significantly different at a $p < 0.05$ level when compared to [27, 214]. Note that for methods [36, 184, 190, 272] a comparison is not possible as the data is not publicly available.

Methods	EDV (mL)		ESV (mL)		EF(%)	
	R	LOA	R	LOA	R	LOA
BEAS	0.953	-3.29 ± 19.03	0.960	-4.84 ± 16.09	0.911	1.7 ± 5.18
Queirós <i>et al.</i> [214]	0.965	-4.99 ± 17.66	0.971	-5.83 ± 13.14	0.927	2.30 ± 4.20
Barbosa <i>et al.</i> [27]	0.965	-4.99 ± 17.66	0.967	-6.78 ± 13.86	0.889	2.88 ± 5.24
Bernier <i>et al.</i> [36]	0.979	2.74 ± 13.87	0.968	2.18 ± 13.73	0.811	0.05 ± 7.84
van Stralen <i>et al.</i> [272]	0.958	-4.86 ± 18.08	0.965	-15.39 ± 15.08	0.751	8.40 ± 7.72
Oktay <i>et al.</i> [184]	0.961	-4.14 ± 17.35	0.973	-3.47 ± 13.62	0.892	0.48 ± 5.50
Inter-observer variability <i>et al.</i> [190]	0.981	-0.64 ± 9.27	0.987	-0.50 ± 7.35	0.959	0.13 ± 3.10

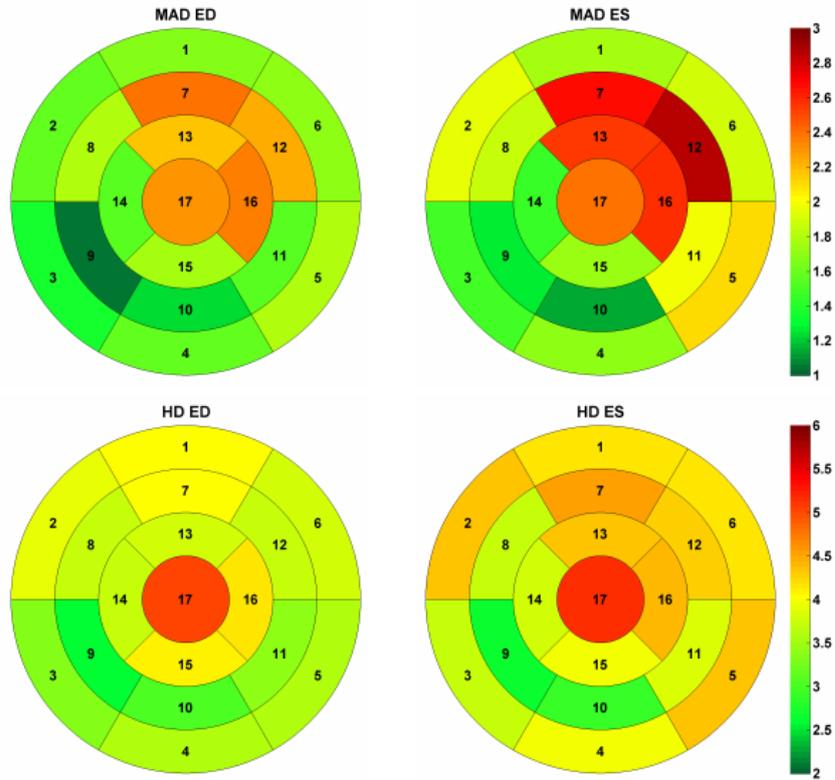


FIGURE 10 – Bullseye plots of average MAD and HD at every region for ED and ES for the training datasets. Measures in mm.

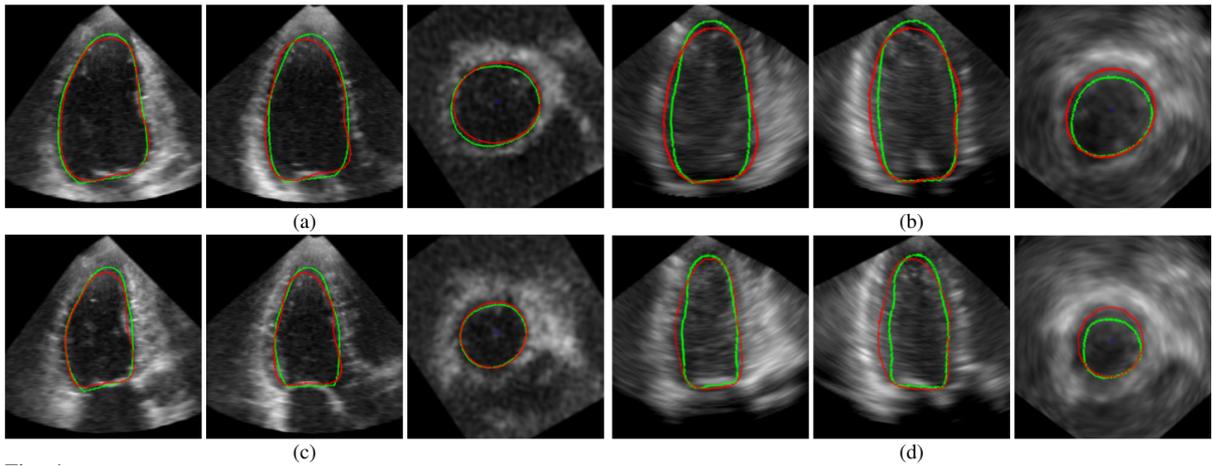


FIGURE 11 – Best (a,c) and worst (b,d) automatic segmentation results (red) compared to manual contours by experts (green) at ED (top row) and ES (bottom row) from the CETUS training set. The three orthogonal planes shown for each 3D image were chosen according to the automatically defined LAx/SAx orientation.

Regarding computational time, the proposed framework took on average 0.9 s for the initialization, 0.6 s for the SAX orientation and a combined time of 1.1 s for the two stages of ED segmentation. The tracking took on average 0.8 s/frame and the final ES segmentation 0.4 s. The total time for a fully automatic ED/ES segmentation was on average 11 s. All data was processed in a non-optimized MATLAB implementation.

4.5.2 Position/orientation performance and sensitivity

At initialization, the position and LAX orientation errors were respectively 3.7 ± 2.1 mm and $5.0 \pm 2.8^\circ$. After refinement at the first stage of ED BEAS segmentation, the position and LAX orientation errors were reduced to respectively 2.4 ± 1.0 mm and $4.4 \pm 2.4^\circ$. Automatic SAX orientation failed in one of the cases due to low image quality giving an error of 120.2° compared to manual annotation of the RV insertion point. On the remaining datasets the SAX orientation error was $6.9 \pm 4.4^\circ$.

Fig. 12 shows the influence on the segmentation performance of the position and orientation of the automatically defined BEAS coordinate system with respect to the position and orientation of the reference meshes. It can be observed that the position and LAX orientation have the most influence on the segmentation results, where a distance above 2mm from the reference mesh centroid or an LAX angle deviation greater than 8° give an error larger than what was obtained with the fully automatic method used in this study.

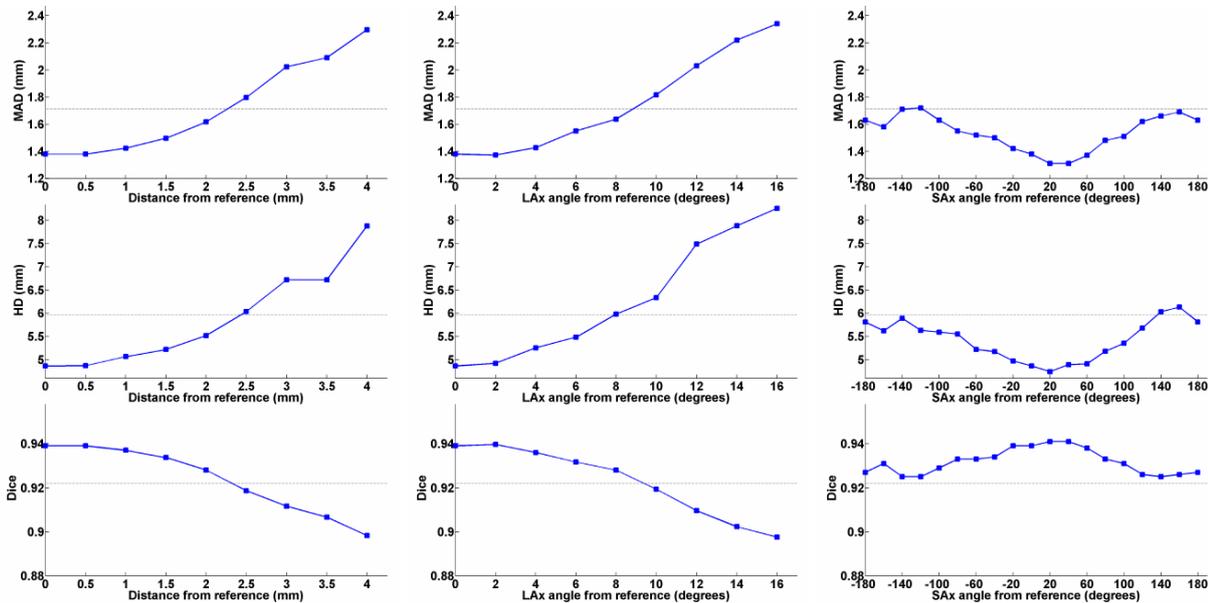


FIGURE 12 – Influence of the distance and angle error from the reference position and orientation on the distance metrics (MAD, HD and Dice) at ED. Horizontal dotted line indicates the performance obtained with the proposed automatic framework on the CETUS training set.

4.5.3 Parameter sensitivity assessment

Fig. 13 shows the influence of the parameters α , β and γ on the segmentation results at ED and ES. For the interval considered from 50% to 150% of the preset value, none of the observed changes were statistically significant at a $p < 0.01$ level and only the MAD at ES showed several statistically significant changes at a $p < 0.05$ level when changing β . When parameters β and γ are set to 0, the difference is statistically significant at a $p < 0.001$ level whereas for α the difference is not statistically significant. When removing the IAAOF, the ES segmentation presents an MAD, HD and Dice of $2.9 \pm 1.1\text{mm}$, $9.8 \pm 2.9\text{mm}$ and 0.861 ± 0.054 respectively (all statistically significant at a $p < 0.001$ level).

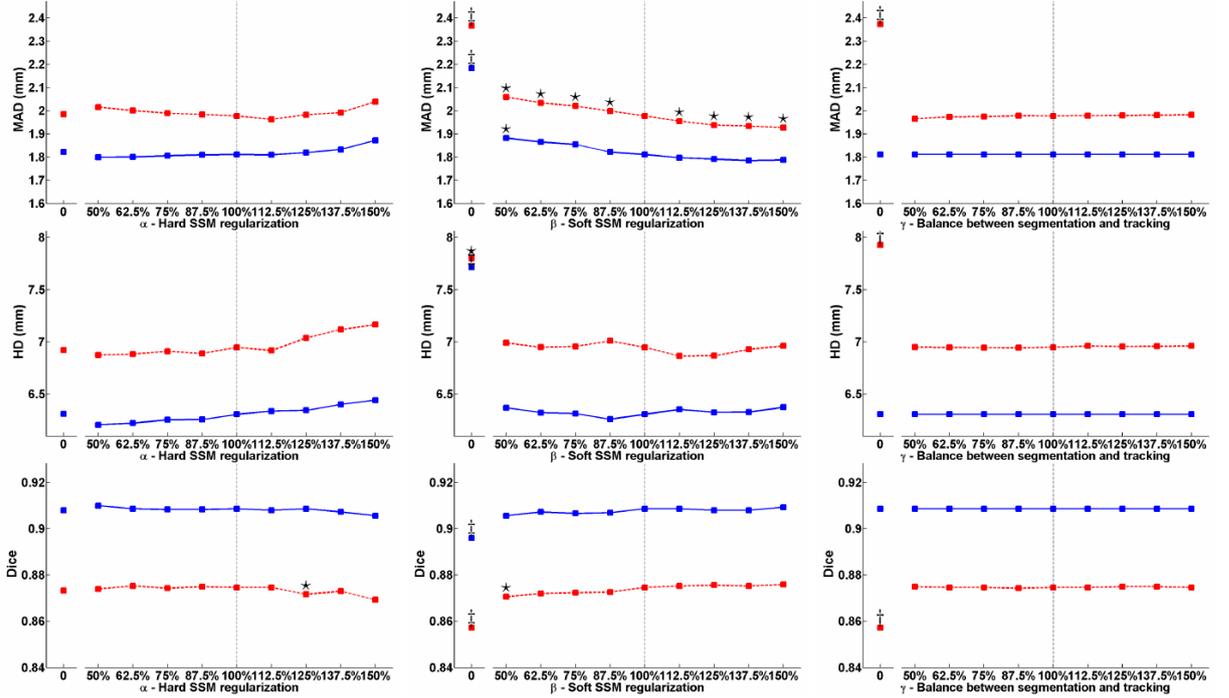


FIGURE 13 – Influence of the variation of each of the considered parameters α , β and γ on the distance metrics (MAD, HD and Dice) at ED (blue) and ES (red dotted). Vertical dotted line indicates the preset parameter value. \star , \dagger and \ddagger indicate respectively that the difference to the result with the preset values was statistically significant at a $p < 0.05$, $p < 0.01$ and $p < 0.001$ level.

4.6 Discussion

A fully automatic LV segmentation and tracking framework is proposed, combining the strengths of image information from BEAS and shape-based clues from an SSM for segmentation and IAAOF to perform tracking. The way in which the SSM is represented on the BEAS space, through the corresponding B-spline representation coefficients $c[\mathbf{k}]$, brings BEAS and the SSM closer together, avoiding steps such as conversion between the spherical and Cartesian coordinate systems and scaling/translation operations. It also avoids one of the fundamental problems with SSM, the point correspondence between different training shapes and with testing shapes. This approach assumes however that the position and orientation of the coordinate system is identical

for every shape. For the training shapes, it is trivial to match the position and orientation of every shape, making the previous assumption valid. When trying to fit the SSM to a new image, the center and both the LAx and SAx orientations have to be guessed from image features.

4.6.1 Segmentation performance

From Table 5 it is clear that the proposed automatic method shows excellent segmentation and tracking performance and outperforms any other of the state-of-the-art approaches applied to the same database. Compared to other approaches using BEAS [27, 214], the impact of the SSM regularization on ED segmentation is statistically significant. With a better starting point at ED for the IAAOF, together with the SSM regularization at ES, the ES segmentation results are also improved, thereby outperforming other state-of-the-art methods. Given the different strategies used in each framework, it is difficult to say with certainty what is the reason behind the differences in performance but the following possible reasons can be considered : regarding the semi-automatic method of Bernier *et al.* [36] using graph cuts, this method lacks a source of prior information needed to give an accurate segmentation when image information is low or incongruous. For both van Stralen *et al.* [272] and Oktay *et al.* [184] that information is provided, respectively, by an active appearance model and a multi-atlas approach. However, both these approaches use ultrasound data as a prior which can be more variable than cMR, especially for reduced datasets. Moreover, both these methods intend to model the appearance of the image, which can be particularly difficult due to the differences between vendors, bad acquisition window or the presence of artifacts. Regarding the clinical indices on Table 6, the proposed method has a performance similar to the remaining state-of-the-art methods.

Regarding the regionwise analysis shown in Fig. 10, there could be two possible explanations for the regions with larger error : either there are inherent image characteristics that make segmentation more difficult or there are framework specific characteristics that cause these errors, such as a systematic error on the LAx detection. However, regionwise error analysis in different frameworks and on manual contouring by experts replicate this trend of larger errors at the apical and anterolateral regions [30], which points to inherent image characteristics that make the segmentation more difficult. Indeed, at the apex, image information is low due to noise in the near field, whereas for the anterolateral region, dropout in this region is common due to its position and proximity to lung tissue.

As for the computational speed, the proposed framework continues to be computationally efficient, especially if compared to other state-of-the-art approaches. Oktay *et al.* [184] reported an average time of 16 min per image and Van Stralen *et al.* [274] reported an average segmentation time of 15s in a C++ environment [271] to which the tracking time must be added (not reported). Furthermore, one can consider ways of decreasing the computational burden of the proposed method by changing to a more efficient implementation in C++, where it has been shown that 3D endocardial segmentation can be done using BEAS in approximately 12.5 ms [25].

4.6.2 Position/orientation performance and sensitivity

As predicted, moving the position and orientation away from the reference has a strong impact on the performance. The fact that SAx orientation has a smaller effect than center position and LAx orientation can be explained by the fact that, though the LV is far from being symmetric, the shape differences between the different sides are much less pronounced than the

shape difference between the apex and base of the LV or those resulting from representing the LV shape from a wrong position. As such, a compromise between the image information and the SSM can more easily be found for an incorrect SAx orientation than from an incorrect center position or LAx orientation.

Fig. 12 also shows that one of the bottlenecks of this method is the positioning and orientation of the LV. It can be seen that when the reference position and orientation is used, the error decreases considerably (MAD : 1.4 mm ; HD : 4.9 mm ; Dice : 0.959). As such, it would be important, in future work, to focus on better automatic initialization methods that, ideally, would provide the true center of the LV and the LAx and SAx orientation. This would imply however to move away from the current initialization, which roughly delineates the LV using the Hough transform, to more complex methodologies, possibly involving machine learning or other more abstract approaches.

4.6.3 Parameter sensitivity assessment

Overall, the parameter sensitivity assessment showed that the performance of the proposed method is not significantly impaired within a wide range of the parameter settings. The parameters related to the SSM regularization seem to have a higher impact as they control the balance between the image information and the SSM. The parameter related to the balance between segmentation and tracking has, as expected, no impact on ED segmentation since γ is not used at ED, and little impact on ES segmentation performance. When each of the parameters is set to zero, thus turning off the corresponding energy contribution, the performance contribution of each energy becomes clear and both β and γ are crucial for the results obtained. The contribution of α is, however, less pronounced. This is due to the fact that the soft energy term already penalizes shapes away from the mean shape, making it less likely for the segmented shape to deviate to the hard set limits at $m=2.5$. Nevertheless, it can be argued that the hard energy term is important to effectively limit the maximum deviation from the mean shape (if $\alpha = 1$) and in more challenging images where image artifacts could make it easier for the segmented shape to deviate from the mean.

Regarding the IAAOF, it is shown that it also plays an important role in following the endocardial surface from ED to ES to initialize the segmentation at ES, as the results without the IAAOF are significantly worse than the proposed method. Nevertheless, in spite of the fact that in this study the IAAOF was chosen to track the endocardial surface, other tracking methods could equally be applied in a straightforward manner and, if proven to be more effective in tracking the LV, could potentially improve the ES segmentation results further. Though in this study only the parameters related to the balance of the different energies were studied, the performance of BEAS and the IAAOF also depend on different parameters. Nonetheless, these have been studied before [22, 196, 214] and the optimal settings found were used in this study.

4.6.4 Limitations and future work

In spite of the promising results shown in this paper, there are limitations which must be addressed in the future. First, as mentioned in Section 4.6.2, the positioning and orientation of the LV is a limiting factor of the accuracy of the proposed framework and should be addressed in the future to provide better segmentation results. Secondly, the parameter tuning performed in this study was quite limited. While in this study only parameters β and γ were subject to parameter tuning, there are other parameters that could be further tuned and which were not

directly addressed. Even though some of these have been tuned before on the same dataset such as the BEAS [27] and IAAOF [214] parameters, a tuning of all parameters together could prove beneficial, especially for the framework elements identified as crucial such as the initialization. Thirdly, in this study only the endocardial border was considered. Nevertheless, the epicardial border is also of importance to study clinical indices, such as LV mass, and is an essential step for automatic cardiac strain measurements through the definition of a region-of-interest. As such, it would be interesting to build an SSM that would describe both the endo- and epicardial borders so that the current framework could be applied for full myocardial segmentation. However, the validation of such a framework cannot be done with the CETUS challenge dataset, as no epicardial contours are provided and, to the author's knowledge, there are no other freely available and reliable datasets of 3D ultrasound data with both endo- and epicardial manual contours.

The dataset used for the SSM must also be considered. First, it could be argued that the cMR shapes used are not ideal as they are derived from 2D slices rather than from true 3D data. However, that would imply that replacing the current SSM by one built from true 3D data would only further improve the results as more accurate data would be embedded into the SSM. Secondly, the very population targeted by the study from where the shapes were obtained is not ideal. Given that DOPPLER-CIP targeted patients suspected of chronic ischemic disease, one cannot consider that the dataset used represents a normal population. However, as before, that would imply that replacing this population with a more representative one would only improve results as the SSM is more well suited for the purpose for which it is intended.

5 3rd contribution : deep convolutional neural network study on RT2DE

In this section we present a study that we put in place to evaluate how far the state-of-the-art deep convolutional network methods can go at assessing 2D echocardiographic images, *i.e.* segmenting cardiac structures as well as estimating clinical indices. Comparisons were made with non-deep learning state-of-the-art methods which obtained among the best results on the CETUS database [30]. Results were obtained through the CAMUS database described in Sec. 3.1. This work has been the subject of several publications [137, 138, 245].

5.1 Motivations

The lack of publicly-available large scale dataset in 2D echocardiography has prevented a thorough evaluation of the potential of deep learning methods to estimate clinical indices, while these techniques are actively applied with great success for other modalities [33]. Indeed, while the number of medical imaging challenges comparing deep learning methods has exploded this last decade, only one focused on cardiac ultrasound image segmentation [30]. Unfortunately, since the challenge was held in 2014, none of the participant used convolutional neural networks (CNNs) because these methods had not yet gained popularity in medical imaging. The lack of well-annotated echocardiographic datasets can be explained by the difficulty of exporting data from clinical ultrasound equipments and getting a large amount of images carefully annotated by cardiologists due to the very nature of echocardiography. In this context, thanks to the CAMUS dataset presented in Sec. 3.1, the purpose of this study is to provide answers to the following four questions :

1. How well do CNNs perform compared to non-deep learning state-of-the-art techniques?
2. How many patients are needed to train a CNN to get highly accurate results in 2D echocardiographic image segmentation?
3. How accurate can the volumes and ejection fraction be estimated from the segmentation of CNNs compared to the inter/intra-expert variability?
4. What improvement can be brought by sophisticated architectures compared to simpler CNN models for 2D echocardiographic segmentation?

5.2 Evaluated methods

5.2.1 CNN techniques based on an encoder-decoder architecture

The goal of this study is to assess how far CNNs can go at segmenting 2D echocardiographic images. As such, we chose to focus on the well-known encoder-decoder networks (EDNs) which have been the cornerstone of a wide variety of CNNs that were successfully applied in medical imaging [147]. EDNs are based on a two-stage convolutional network architecture well suited for segmentation. The first part, known as the encoder, consists of a series of convolutions and downsampling operations. These operations extract features from the images while spatially compressing them, thus enabling extraction of high-level features. The second part is the decoder,

which uses features from the encoder and applies a set of convolutions and upsampling operations to gradually transform feature maps into a final segmentation.

Among the existing EDNs, one of the most popular architectures used in medical imaging corresponds the U-Net model proposed by Ronneberger *et al.* in 2015 [225]. This network integrates residual connections between the encoder and decoder parts with the goal of helping in retrieving details that were potentially lost during the downsampling while also stabilizing gradients. The original U-Net follows a specific scheme of convolutions, where each downsampling and upsampling step is preceded by two 3x3 convolutional layers, while the number of features is doubled per downsampling and reduced in half per upsampling. U-Net has been successfully applied to a wide range of medical applications [147], but for each application, the network design has usually been adapted and optimized to get the best segmentation performance on each application. The main U-Net design choices can be classified in three categories : *i)* layer choices : convolutional layer size, activation functions, normalization layers, down- and upsampling strategies (*e.g.* max pooling, striding, deconvolution and repeat); *ii)* the optimization process (gradient descent strategy, weight initialization, loss function, batch size, regularization constraints, stopping criteria, deep supervision, dropout); *iii)* data handling (pre-processing, augmentation, sampling). Since the seminal paper in 2015, several studies based on the EDN structure have been carried out with the goal of outperforming the U-Net. Among those methods, two types of approaches have been proposed : those based on U-Net architecture but with extensions such as shape regularization [183] and those with more sophisticated architectures [174, 302]. In this context, we decided to benchmark the following EDNs for the purpose of segmenting 2D echocardiographic images :

5.2.2 U-Net

Taking into account the wide range of possible U-Net designs, we decided to compare the performance of two independent implementations, *i.e.* U-Net 1 optimized for speed, and U-Net 2 optimized for accuracy. This leads to two different architectures (which both differ from the original one proposed by Ronneberger *et al.*) with their own hyperparameters settings, as shown in table 7 and table 8. U-Net 1 & 2 enable to investigate the impact of hyperparameters choices on the quality of the results.

5.2.3 ACNN

Starting from a given segmentation architecture, this method integrates an auxiliary loss to constrain the segmentation output to fit a non-linear compact representation of the underlying anatomy derived from an auto-encoder network [183]. For comparison purposes, we used the U-Net 1 architecture described in table 7 as the segmentation module in our ACNN implementation. Moreover, the following choices were made to obtain the best results on our dataset : *i)* a code of 32 coefficients was set for the auto-encoder network (which allows an average reconstruction accuracy of 97%); *ii)* the hyperparameter balancing the segmentation and shape regularization losses was set so that the two losses had close initialization values.

5.2.4 SHG

Stacked Hourglasses (SHG) method integrates three successive encoder-decoder networks (usually three times the same architecture) where the first two are used as residual blocks [174]. Each output of the encoder-decoder networks is associated with an intermediate segmentation

loss. This strategy is named as deep supervision. The output of the third network is used as the final segmentation result. For comparison purposes, we also used the U-Net 1 architecture as the key encoder-decoder network in our SHG implementation.

5.2.5 U-Net++

This method is also based on deep supervision technique but with the integration of additional convolution layers in the form of dense skip connections [302]. Starting from the official online version of the code, we adapted the corresponding architecture to obtain the best results on our dataset. The following changes were made : *i*) dropout was removed ; *ii*) averaging of the last feature maps of the intermediate outputs was removed ; *iii*) the original design of layers was adapted according to the choices we made to optimized U-Net 1 architecture ; *iv*) the batch size was set to 20.

Please note that the same data pre- and post-processing strategies were applied for each of the five tested EDNs.

5.2.6 Non-deep learning state-of-the-art techniques

To compare the performance of the EDN methods described above, we implemented the following non-deep learning state-of-the-art methods which obtained among the best results during the CETUS challenge [30] and which were recently improved [197] and applied in 2D [137].

SRF - Structured Random Forests (SRF) refer to an ensemble learning method for classification or regression. It operates at training time by building a set of decision trees that assign a label patch to each input image patch, computed as the mean prediction of the individual trees [66]. During the training phase, each tree individually learns a set of split functions from a random subset of the training dataset and input features. Those functions are intended to group patches sharing close image intensities and segmentation patterns. During the testing phase, the image to segment is fragmented into different overlapping patches. Each image patch goes through the splitting functions of each tree so that the mean label patch computed from the reached leaves forms its segmentation. Detailed description of the SRF algorithm implemented in this project can be found in [137]. Compared to our previous study, data was not split between ES and ED nor between 4 chambers and 2 chambers views but processed in one indistinctive pool of images. Since CAMUS has a larger number of patients than the dataset used in [137], we trained 12 individual trees for each subset of 100 patients.

BEASM - The key concept of the B-Spline Explicit Active Surface Model (BEASM) framework is to consider the boundary of a deformable interface as an explicit function, where one of the coordinates of the points within the surface is given explicitly as a function ψ of the remaining coordinates (see Section 4 for more details on this method). In this framework, ψ is defined as a linear combination of B-spline basis functions whose controlled knots are located on a regular rectangular grid defined on the chosen coordinate system (polar space in our case). Based on a standard variational approach, the evolution of the deformable surface is then governed by the minimization of an energy function according to the B-spline coefficients [25]. This framework has been successfully applied in [196] for the coupled segmentation of the LV_{Endo} and LV_{Epi} structures in echocardiography and further extended by the integration of a shape prior directly into the B-spline space in [197], named as BEASM in the rest of the paper. Because BEASM amounts to a deformable-based model, the initialization of the contour plays a crucial role on the quality

TABLE 8 – Main characteristics of the U-Net 2 implementation

U-NET 2 ARCHITECTURE				
Level	Layer	Kernel / Pool size	Activation	Connection
D1	Conv	48 (3,3)		
	BatchNorm		ReLU	
	Conv	48 (3,3)		
	BatchNorm		ReLU	*
	MaxPooling	(2*2)		
D2	Conv	96 (3,3)		
	BatchNorm		ReLU	
	Conv	96 (3,3)		
	BatchNorm		ReLU	**
	MaxPooling	(2*2)		
D3	Conv	192 (3,3)		
	BatchNorm		ReLU	
	Conv	192 (3,3)		
	BatchNorm		ReLU	***
	MaxPooling	(2*2)		
D4	Conv	384 (3,3)		
	BatchNorm		ReLU	
	Conv	384 (3,3)		
	BatchNorm		ReLU	****
	MaxPooling	(2*2)		
D5	Conv	768 (3,3)		
	BatchNorm		ReLU	
	Conv	768 (3,3)		
	BatchNorm		ReLU	
U1	ConvTranspose	384 (2,2) - s(2,2)		
	BatchNorm		ReLU	
	Conv	384 (3,3)		****
	BatchNorm		ReLU	
	Conv	384 (3,3)		
	BatchNorm		ReLU	
U2	ConvTranspose	192 (2,2) - s(2,2)		
	BatchNorm		ReLU	
	Conv	192 (3,3)		***
	BatchNorm		ReLU	
	Conv	192 (3,3)		
	BatchNorm		ReLU	
U3	ConvTranspose	96 (2,2) - s(2,2)		
	BatchNorm		ReLU	
	Conv	96 (3,3)		**
	BatchNorm		ReLU	
	Conv	96 (3,3)		
	BatchNorm		ReLU	
U4	ConvTranspose	48 (2,2) - s(2,2)		
	BatchNorm		ReLU	
	Conv	48 (3,3)		*
	BatchNorm		ReLU	
	Conv	48 (3,3)		
	BatchNorm		ReLU	
Seg	Conv	4 (1,1)	Softmax	

s : height and width strides

of the results. We thus decided to implement two different strategies : *i*) one named BEASM-fully where the evolving contour is automatically initialized from a method inspired by the work proposed in [22] ; *ii*) another named BEASM-semi where the evolving contour is initialized from three points (two at the base and one at the apex of the LV_{Endo} structure) extracted from the reference contours. By doing so, we gave the possibility to quantify the influence of the initialization procedure for BEASM on an heterogeneous ultrasound dataset.

5.3 Results

As stated in Sec. 3.1, the 19% of poor quality images were not used to compute the different metrics provided in this part. Moreover, to avoid the use of different models according to the acquisition settings, we trained only one model for each machine learning method on both apical four-chamber and two-chamber views regardless of the time instant in the cardiac sequence.

5.3.1 Evaluation metrics

Geometrical metrics - To measure the accuracy of the segmentation output (LV_{Endo} , LV_{Epi} or LA) of a given method, the Dice metric, the mean absolute distance (d_m) and the 2D Hausdorff distance (d_H) were used. The Dice similarity index is defined as

$$D = 2(|S_{\text{user}} \cap S_{\text{ref}}|) / (|S_{\text{user}}| + |S_{\text{ref}}|)$$

and is a measure of overlap between the segmented surface S_{user} extracted from a method and the corresponding reference surface S_{ref} . The Dice index gives a value between 0 (no overlap) and 1 (full overlap). d_m corresponds to the average distance between S_{user} and S_{ref} while d_H measures the local maximum distance between the two surfaces.

Clinical metrics - We gauge the methods' performance with 3 clinical indices : *i*) the ED volume (LV_{EDV} in ml) ; *ii*) the ES volume (LV_{ESV} in ml) ; *iii*) the ejection fraction (LV_{EF} as a percentage) for which we computed four metrics : the correlation (*corr*), the bias and the standard deviation (*std*) values (computed from conventional definitions) and the mean absolute error (*mae*). The combination of the bias and standard deviation also provides useful information on the corresponding limit of agreement values.

5.3.2 Empirical results

Geometrical scores : Table 9 shows the segmentation testing accuracy computed from patients having good and medium image quality (406 patients) for the 8 algorithms described in Sec. 5.2. Mean and standard deviation values for each metric were obtained from cross-validating on the 10 folds of the database. The values in bold correspond to the best scores for each metric. From these results, one can see that the EDN implementations get the overall best segmentation scores on all metrics, for both ED and ES. Interestingly, while the EDN methods are fully-automatic, they still get better segmentation results than the semi-automatic BEASM algorithm.

The two U-Nets achieve equivalent results for all the metrics compared to the ones obtained by the more sophisticated encoder-decoder architectures. This hints to the idea that a plateau has been reached, which classical tuning, shape regularization techniques and more sophisticated architectures have difficulties to overcome. This also suggests that a U-Net implementation, which requires less parameters than SHG and U-Net++ methods and less training time than

ACNN, offers the best compromise between the network size and performance for the particular task of 2D echocardiographic image segmentation.

To assess the influence of the layer design in the performance between U-Net 1 and 2, statistical significance of their respective results was analyzed by performing the Wilcoxon signed-rank test for each metric. Based on table 9, results from the two U-Nets are statistically different (p-value < 0.05) for most metrics, apart for the LV_{Epi} Hausdorff distance. However, this must be qualified by the fact that *i*) the U-Net geometrical scores are very close (mean d_m and d_H difference of 0.1 mm and 0.1 mm, respectively), producing distributions with high degree of overlap as shown in Fig. 14; *ii*) the U-Net geometrical results lie between the inter-observer and intra-observer scores for all metrics, proving the robustness of this method in obtaining accurate segmentation results.

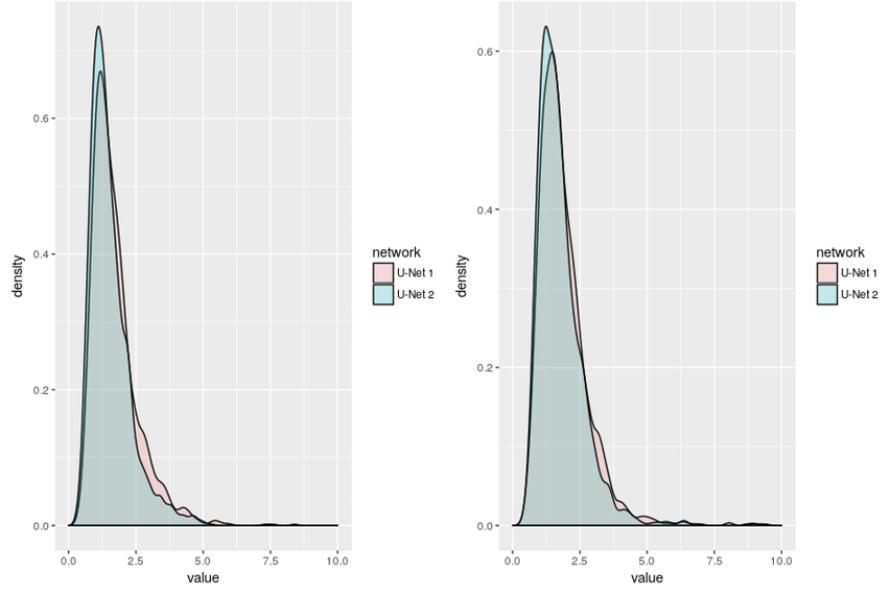
As for the fully automatic non-deep learning state-of-the-art methods, BEASM-fully obtained on average better Hausdorff distances (mean d_H of 9.9 mm at ED and 10.5 mm at ES) while the SRF got better Dice and d_m scores (mean d_m of 3.0 mm at ED and 3.5 mm at ES). However, the large standard deviation values for the SRF illustrate the difficulties of this method in obtaining consistent segmentations over the entire dataset. As for the BEASM-semi, one can see that the manual initialization has a strong impact on the quality of the results, with a mean improvement of 0.8 mm and 2.4 mm for the d_m and d_H metrics, respectively. Moreover, it is well known that the left ventricle shape is more difficult to segment at ES, leading to slightly worse performance for classical algorithms on this time instant. This property is also confirmed in our study since all the evaluated methods produced better results at ED on every metric.

As complement, we provided in table 10 the geometrical scores obtained on the poor quality images (94 patients) for the 8 evaluated algorithms. For this part of the dataset, the EDNs also obtained the best segmentation results on all metrics. Interestingly, while EDN scores on poor quality images are slightly worse than those computed on good and medium quality, they remain very competitive compared to the scores given in table 9 (mean LV_{Endo} d_m and d_H of 2.2 mm and 7.0 mm and mean LV_{Epi} d_m and d_H of 2.3 mm and 7.6 mm). To allow visual assessment of the segmentation performance of the different methods implemented in this work, we provide in Fig. 15 to 21 the segmentation results obtained by each of the presented methods and the cardiologists on a given patient with a good image quality.

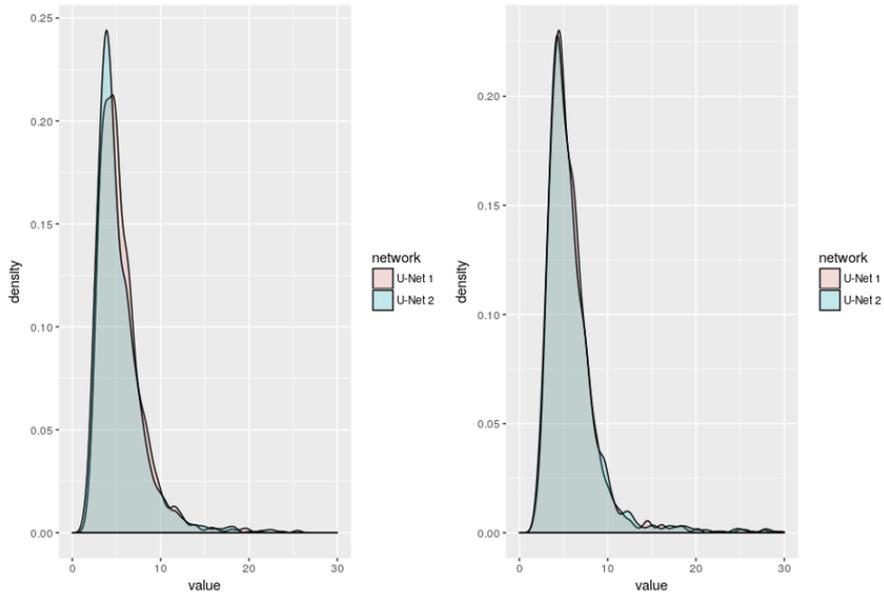
TABLE 9 – Segmentation accuracy (LV_{Endo} and LV_{Epi}) of 8 evaluated methods on the ten test datasets restricted to patients having **good and medium image quality** (406 patients in total). The values in bold refer to the best performance for each measure. p-values are based on the Wilcoxon signed rank between U-Net 1 and 2 for each evaluation metric. The inter and intra-observer measures were all computed from fold 5 restricted to patients having good & medium image quality (40 patients)

Methods *	ED						ES					
	LV_{Endo}			LV_{Epi}			LV_{Endo}			LV_{Epi}		
	D	d_m	d_H									
	val.	mm	mm									
O_{1a} vs O_2 (inter-obs)	0.919 ± 0.033	2.2 ± 0.9	6.0 ± 2.0	0.913 ± 0.037	3.5 ± 1.7	8.0 ± 2.9	0.873 ± 0.060	2.7 ± 1.2	6.6 ± 2.4	0.890 ± 0.047	3.9 ± 1.8	8.6 ± 3.3
O_{1a} vs O_3 (inter-obs)	0.886 ± 0.050	3.3 ± 1.5	8.2 ± 2.5	0.943 ± 0.018	2.3 ± 0.8	6.5 ± 2.6	0.823 ± 0.091	4.0 ± 2.0	8.8 ± 3.5	0.931 ± 0.025	2.4 ± 1.0	6.4 ± 2.4
O_2 vs O_3 (inter-obs)	0.921 ± 0.037	2.3 ± 1.2	6.3 ± 2.5	0.922 ± 0.036	3.0 ± 1.5	7.4 ± 3.0	0.888 ± 0.058	2.6 ± 1.3	6.9 ± 2.9	0.885 ± 0.054	3.9 ± 1.9	8.4 ± 2.8
O_{1a} vs O_{1b} (intra-obs)	0.945 ± 0.019	1.4 ± 0.5	4.6 ± 1.8	0.957 ± 0.019	1.7 ± 0.9	5.0 ± 2.3	0.930 ± 0.031	1.3 ± 0.5	4.5 ± 1.8	0.951 ± 0.021	1.7 ± 0.8	5.0 ± 2.1
SRF	0.895 ± 0.074	2.8 ± 3.6	11.2 ± 10.2	0.914 ± 0.057	3.2 ± 2.0	13.0 ± 9.1	0.848 ± 0.137	3.6 ± 7.8	11.6 ± 13.6	0.901 ± 0.078	3.5 ± 4.7	13.0 ± 11.1
BEASM-fully	0.879 ± 0.065	3.3 ± 1.8	9.2 ± 4.9	0.895 ± 0.051	3.9 ± 2.1	10.6 ± 5.1	0.826 ± 0.092	3.8 ± 2.1	9.9 ± 5.1	0.880 ± 0.054	4.2 ± 2.0	11.2 ± 5.1
BEASM-semi	0.920 ± 0.039	2.2 ± 1.2	6.0 ± 2.4	0.917 ± 0.038	3.2 ± 1.6	8.2 ± 3.0	0.861 ± 0.070	3.1 ± 1.6	7.7 ± 3.2	0.900 ± 0.042	3.5 ± 1.7	9.2 ± 3.4
U-Net 1	0.934 ± 0.042	1.7 ± 1.0	5.5 ± 2.9	0.951 ± 0.024	1.9 ± 0.9	5.9 ± 3.4	0.905 ± 0.063	1.8 ± 1.3	5.7 ± 3.7	0.943 ± 0.035	2.0 ± 1.2	6.1 ± 4.1
U-Net 2	0.939 ± 0.043	1.6 ± 1.3	5.3 ± 3.6	0.954 ± 0.023	1.7 ± 0.9	6.0 ± 3.4	0.916 ± 0.061	1.6 ± 1.6	5.5 ± 3.8	0.945 ± 0.039	1.9 ± 1.2	6.1 ± 4.6
ACNN	0.932 ± 0.034	1.7 ± 0.9	5.8 ± 3.1	0.950 ± 0.026	1.9 ± 1.1	6.4 ± 4.1	0.903 ± 0.059	1.9 ± 1.1	6.0 ± 3.9	0.942 ± 0.034	2.0 ± 1.2	6.3 ± 4.2
SHG	0.934 ± 0.034	1.7 ± 0.9	5.6 ± 2.8	0.951 ± 0.023	1.9 ± 1.0	5.7 ± 3.3	0.906 ± 0.057	1.8 ± 1.1	5.8 ± 3.8	0.944 ± 0.034	2.0 ± 1.2	6.0 ± 4.3
U-Net ++	0.927 ± 0.046	1.8 ± 1.1	6.5 ± 3.9	0.945 ± 0.026	2.1 ± 1.0	7.2 ± 4.5	0.904 ± 0.060	1.8 ± 1.0	6.3 ± 4.2	0.939 ± 0.034	2.1 ± 1.1	7.1 ± 5.1
p-values	$\ll 0.05$	$\ll 0.05$	$\ll 0.05$	$\ll 0.05$	$\ll 0.05$	≈ 0.83	$\ll 0.05$	$\ll 0.05$	$\ll 0.05$	$\ll 0.05$	$\ll 0.05$	≈ 0.48

* LV_{Endo} : Endocardial contour of the left ventricle ; LV_{Epi} : Epicardial contour of the left ventricle
ED : End diastole ; ES : End systole ; D : Dice index
 d_m : mean absolute distance ; d_H : Hausdorff distance ; mae : mean absolute error



(a) d_m distributions for the LV_{Endo} (left) and LV_{Epi} (right) structures

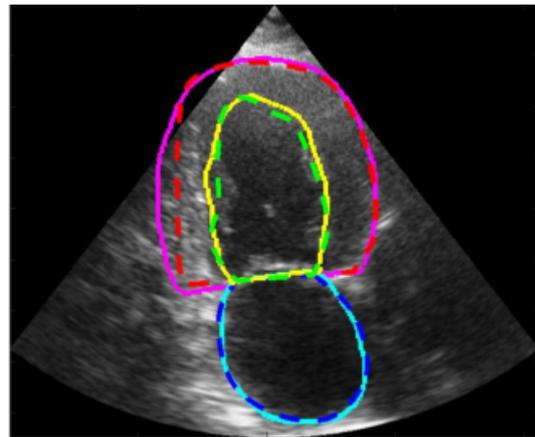


(b) d_H distributions for the LV_{Endo} (left) and LV_{Epi} (right) structures

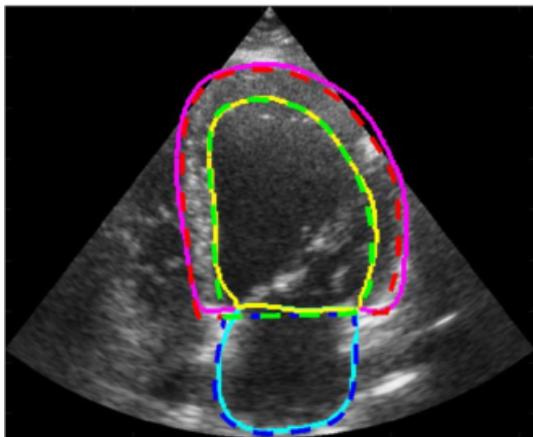
FIGURE 14 – d_m and d_H distributions derived from the U-Net 1 and 2 results obtained on LV_{Endo} and LV_{Epi} structures.



(a) 2CH-ED : d_m -endo = 1.4 , d_m -epi = 1.5 , d_H -endo = 4.4 , d_H -epi = 4.8 mm.



(b) 2CH-ES : d_m -endo = 1.6 , d_m -epi = 2.4 , d_H -endo = 5.1 , d_H -epi = 7.7 mm.

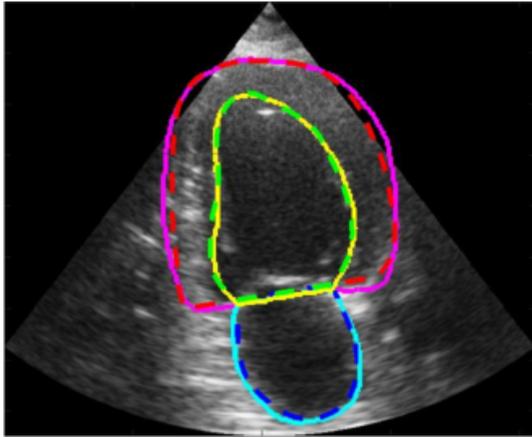


(c) 4CH-ED : d_m -endo = 1.0 , d_m -epi = 2.2 , d_H -endo = 5.9 , d_H -epi = 4.5 mm.

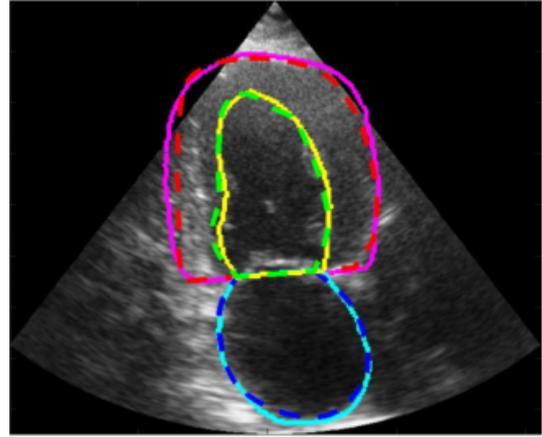


(d) 4CH-ES : d_m -endo = 1.0 , d_m -epi = 1.7 , d_H -endo = 3.8 , d_H -epi = 4.5 mm.

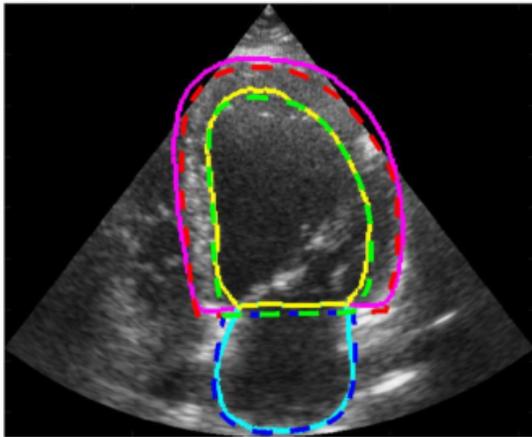
FIGURE 15 – Segmentation results obtained by the U-Net 1 architecture on Patient 27 (image defined as good quality). Ground-truth contours are dotted and prediction contours are drawn in full line.



(a) 2CH-ED : d_m -endo = 1.1 , d_m -epi = 1.4 , d_H -endo = 4.0 , d_H -epi = 4.0 mm.



(b) 2CH-ES : d_m -endo = 1.4 , d_m -epi = 1.2 , d_H -endo = 4.4 , d_H -epi = 4.33 mm.

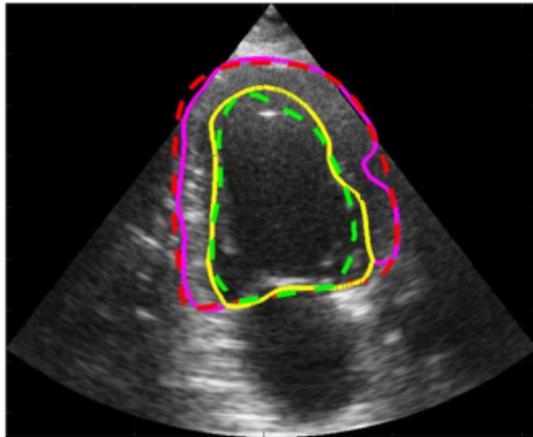


(c) 4CH-ED : d_m -endo = 1.6 , d_m -epi = 2.0 , d_H -endo = 6.8 , d_H -epi = 4.6 mm.

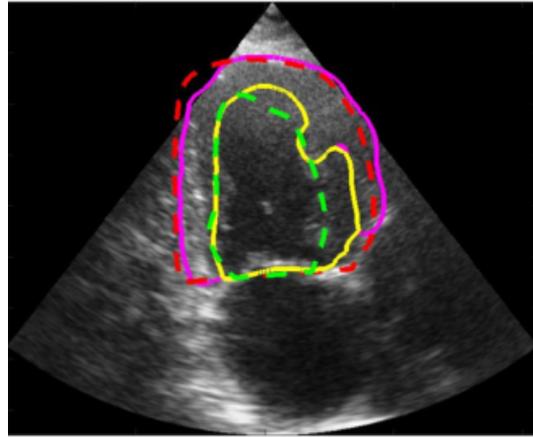


(d) 4CH-ES : d_m -endo = 1.6 , d_m -epi = 1.5 , d_H -endo = 4.7 , d_H -epi = 4.3 mm.

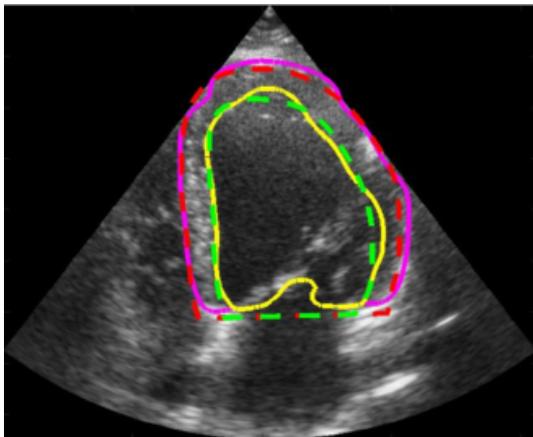
FIGURE 16 – Segmentation results obtained by the U-Net 2 architecture on Patient 27.



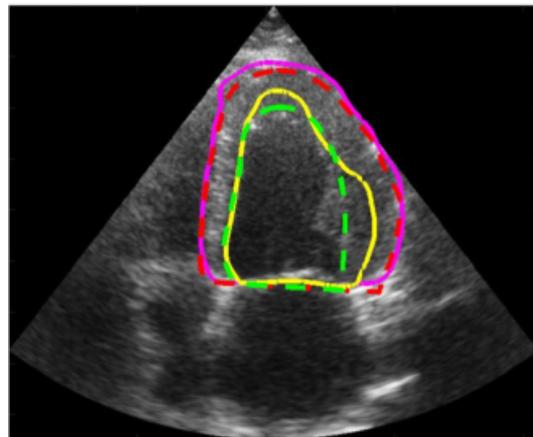
(a) 2CH-ED : d_m -endo = 2.7 , d_m -epi = 1.8 , d_H -endo = 8.5 , d_H -epi = 8.1 mm.



(b) 2CH-ES : d_m -endo = 3.8 , d_m -epi = 2.6 , d_H -endo = 14.0 , d_H -epi = 7.0 mm.

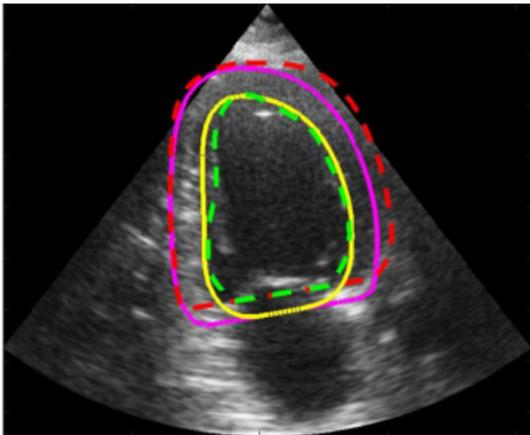


(c) 4CH-ED : d_m -endo = 2.3 , d_m -epi = 2.3 , d_H -endo = 11.1 , d_H -epi = 14.2 mm.

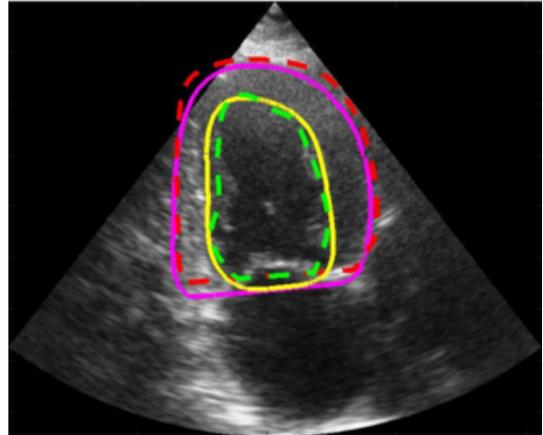


(d) 4CH-ES : d_m -endo = 3.2 , d_m -epi = 1.7 , d_H -endo = 6.8 , d_H -epi = 4.8 mm.

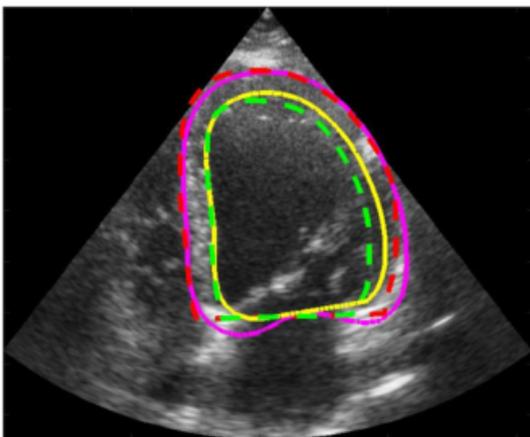
FIGURE 17 – Segmentation results obtained by the SRF method on Patient 27.



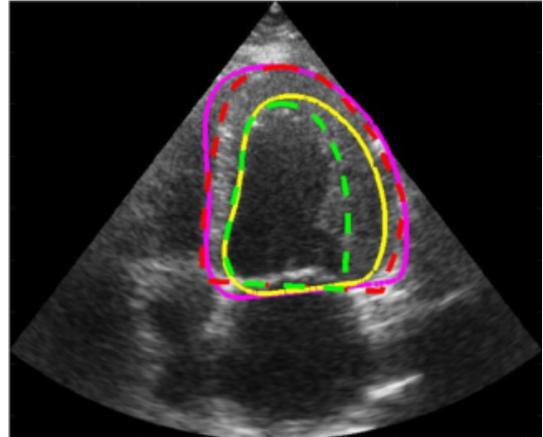
(a) 2CH-ED : d_m -endo = 3.0 , d_m -epi = 3.6 , d_H -endo = 7.2 , d_H -epi = 7.7 mm.



(b) 2CH-ES : d_m -endo = 2.9 , d_m -epi = 2.8 , d_H -endo = 4.7 , d_H -epi = 5.5 mm.

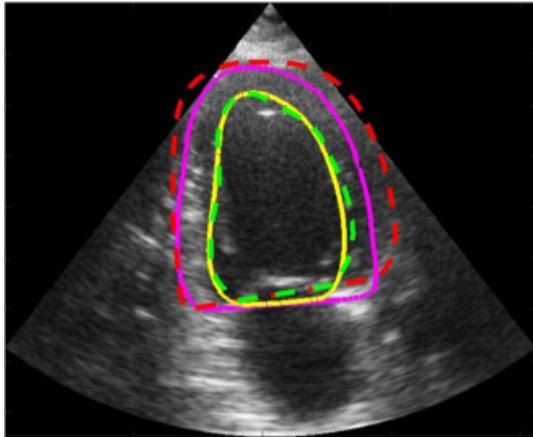


(c) 4CH-ED : d_m -endo = 2.4 , d_m -epi = 1.7 , d_H -endo = 6.2 , d_H -epi = 6.6 mm.

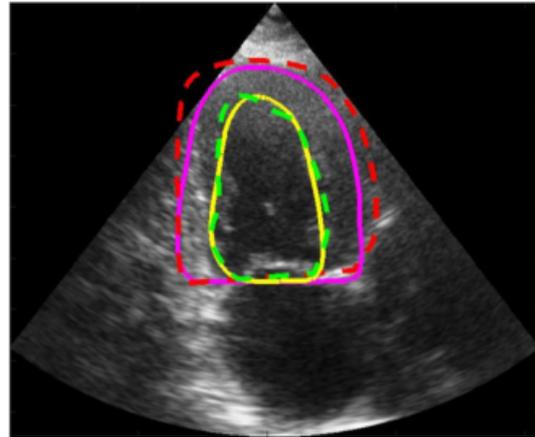


(d) 4CH-ES : d_m -endo = 5.0 , d_m -epi = 2.3 , d_H -endo = 14.2 , d_H -epi = 6.5 mm.

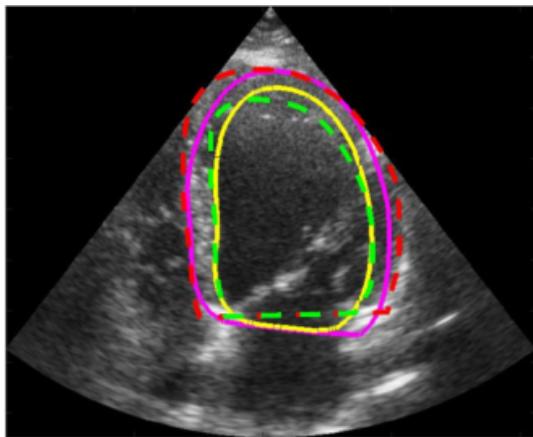
FIGURE 18 – Segmentation results obtained by the BEASM method with automatic initialization on Patient 27.



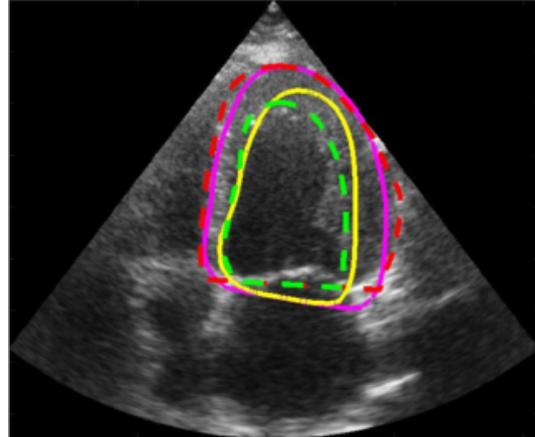
(a) 2CH-ED : d_m -endo = 1.7 , d_m -epi = 4.9 , d_H -endo = 4.1 , d_H -epi = 9.6 mm.



(b) 2CH-ES : d_m -endo = 1.5 , d_m -epi = 3.8 , d_H -endo = 8.5 , d_H -epi = 8.5 mm.



(c) 4CH-ED : d_m -endo = 2.1 , d_m -epi = 3.2 , d_H -endo = 6.2 , d_H -epi = 8.2 mm.



(d) 4CH-ES : d_m -endo = 3.4 , d_m -epi = 2.8 , d_H -endo = 3.5 , d_H -epi = 7.7 mm.

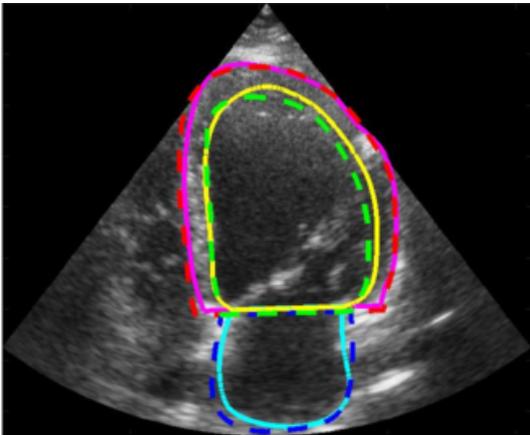
FIGURE 19 – Segmentation results obtained by the BEASM method with semi-automatic initialization on Patient 27.



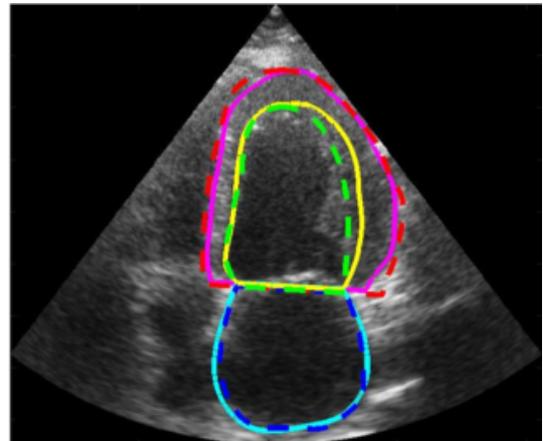
(a) 2CH-ED : $d_{m\text{-endo}} = 3.1$, $d_{m\text{-epi}} = 4.2$, $d_{H\text{-endo}} = 5.9$, $d_{H\text{-epi}} = 9.0$ mm.



(b) 2CH-ES : $d_{m\text{-endo}} = 3.3$, $d_{m\text{-epi}} = 5.5$, $d_{H\text{-endo}} = 8.1$, $d_{H\text{-epi}} = 9.5$ mm.

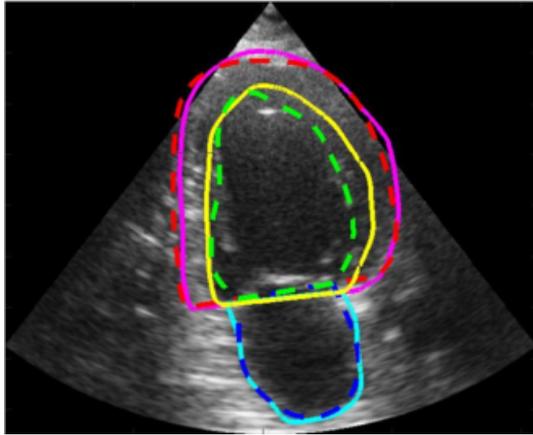


(c) 4CH-ED : $d_{m\text{-endo}} = 2.0$, $d_{m\text{-epi}} = 1.4$, $d_{H\text{-endo}} = 4.8$, $d_{H\text{-epi}} = 3.4$ mm.



(d) 4CH-ES : $d_{m\text{-endo}} = 2.4$, $d_{m\text{-epi}} = 1.9$, $d_{H\text{-endo}} = 6.1$, $d_{H\text{-epi}} = 9.0$ mm.

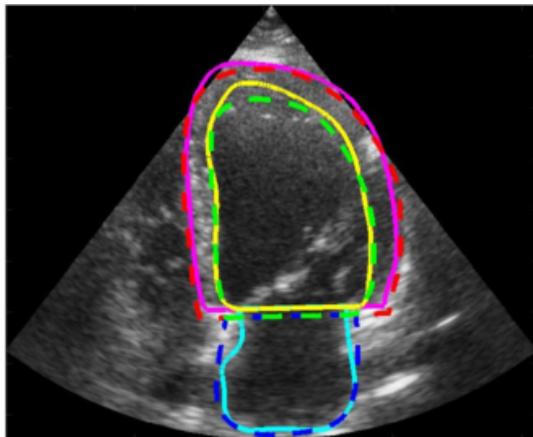
FIGURE 20 – Segmentation results obtained by the cardiologist 2 on Patient 27.



(a) 2CH-ED : d_m -endo = 4.5 , d_m -epi = 4.1 , d_H -endo = 9.7 , d_H -epi = 4.1 mm.



(b) 2CH-ES : d_m -endo = 5.9 , d_m -epi = 4.5 , d_H -endo = 13.2 , d_H -epi = 4.5 mm.



(c) 4CH-ED : d_m -endo = 1.7 , d_m -epi = 4.0 , d_H -endo = 6.8 , d_H -epi = 4.0 mm.



(d) 4CH-ES : d_m -endo = 2.7 , d_m -epi = 4.4 , d_H -endo = 4.9 , d_H -epi = 4.4 mm.

FIGURE 21 – Segmentation results obtained by the cardiologist 3 on Patient 27.

TABLE 10 – LV_{Endo} segmentation accuracy of 5 evaluated methods on the ten test datasets restricted to patients having *poor image quality* (94 patients in total). The values in bold refer to the best performance for each measure. p-values are based on the Wilcoxon signed rank between the U-Net methods for each evaluation metric.

Methods *	ED						ES					
	LV_{Endo}			LV_{Epi}			LV_{Endo}			LV_{Epi}		
	D	d_m	d_H									
	val.	mm	mm									
U-Net 1	0.921 ± 0.037	2.1 ± 1.0	6.5 ± 3.0	0.945 ± 0.021	2.2 ± 1.0	6.8 ± 3.0	0.893 ± 0.059	2.2 ± 1.2	6.8 ± 4.2	0.935 ± 0.031	2.4 ± 1.3	7.2 ± 4.7
U-Net 2	0.921 ± 0.038	2.1 ± 1.0	6.9 ± 3.4	0.947 ± 0.023	2.1 ± 1.0	7.3 ± 4.1	0.899 ± 0.058	2.1 ± 1.2	6.7 ± 3.9	0.937 ± 0.032	2.2 ± 1.2	7.7 ± 4.9
SRF	0.868 ± 0.060	3.7 ± 1.9	14.8 ± 9.5	0.893 ± 0.059	4.2 ± 2.3	16.0 ± 8.4	0.798 ± 0.125	4.7 ± 3.5	17.7 ± 13.3	0.855 ± 0.107	4.9 ± 3.0	18.7 ± 12.1
BEAS fully	0.857 ± 0.083	4.1 ± 2.6	10.9 ± 6.5	0.889 ± 0.057	4.4 ± 2.5	12.0 ± 6.2	0.800 ± 0.101	4.7 ± 2.7	12.3 ± 6.5	0.873 ± 0.062	4.7 ± 2.6	12.5 ± 6.2
BEAS semi	0.915 ± 0.039	2.5 ± 1.3	6.5 ± 2.8	0.914 ± 0.036	3.4 ± 1.6	8.6 ± 3.0	0.860 ± 0.061	3.3 ± 1.5	8.3 ± 3.6	0.900 ± 0.040	3.6 ± 1.6	9.5 ± 3.5
p-values	≈ 0.836	≈ 0.860	≈ 0.441	≈ 0.195	≈ 0.262	≈ 0.724	≈ 0.268	≈ 0.224	≈ 0.839	≈ 0.226	≈ 0.234	≈ 0.735

Clinical scores : Table 11 contains the clinical metrics for the 8 methods. Those indices were computed with the Simpson’s rule [86] from the segmentation results of each algorithm. The values in bold represent the best scores for the corresponding index while the p-values allow to assess the statistical significance between the results of U-Net 1 and 2. As for segmentation, the EDNs obtained the best clinical scores on all the tested metrics (bias was not taken into account since the lowest bias value in itself does not necessarily mean the best performing method). Regarding the estimation of the LV_{EDV} and LV_{ESV} , the EDNs obtained high correlation scores (all above 0.94) and reasonably small biases (at most 11.4 ml), standard deviations (less than 12.9 ml) and mean absolute errors (at most 13.2 ml). Results are more contrasted for the estimation of the LV_{EF} . For this metric, the EDNs got lower correlation scores (at most 0.82) but smaller biases (less than 1.8 %), standard deviations (at most 7.8 %) and mean absolute errors (less than 5.7 %). It is worth pointing out that average EDN scores are all below the inter-observer scores. This proves the clinical interest of such approaches but also reveals the needs for improvement as discussed in Section 5.4. Here again, even if the U-Net methods involved simpler architecture, they obtained similar results compared to the more sophisticated EDNs. Finally, using the Wilcoxon signed-rank test, U-Net 1 and U-Net 2 produced LV_{EDV} , LV_{ESV} and LV_{EF} results whose difference is statistically significant, although their measures are very close.

TABLE 11 – Clinical metrics of the 8 evaluated methods on the ten test folds restricted to patients having *good and medium image quality* (406 patients in total). All the metrics were computed using the annotations of expert O_{1a} . The values in bold correspond to the best performance for the correlation and mean absolute error metrics. p-values are based on the Wilcoxon signed-rank between U-Net 1 and 2 for the LV_{Endo} , LV_{Epi} and LV_{EF} . The inter and intra-observer measures were all computed from fold 5 restricted to patients having good and medium image quality (40 patients).

Methods *	LV_{EDV}			LV_{ESV}			LV_{EF}		
	<i>corr</i>	<i>bias</i> $\pm\sigma$	<i>mae</i>	<i>corr</i>	<i>bias</i> $\pm\sigma$	<i>mae</i>	<i>corr</i>	<i>bias</i> $\pm\sigma$	<i>mae</i>
	val.	ml	ml	val.	ml	ml	val.	%	%
O_{1a} vs O_2 (inter-obs)	0.940	18.7 \pm 12.9	18.7	0.956	18.9 \pm 9.3	18.9	0.801	-9.1 \pm 8.1	10.0
O_{1a} vs O_3 (inter-obs)	0.895	39.0 \pm 18.8	39.0	0.860	35.9 \pm 17.1	35.9	0.646	-12.6 \pm 10.0	13.4
O_2 vs O_3 (inter-obs)	0.926	-20.3 \pm 15.6	21.0	0.916	-17.0 \pm 13.5	17.7	0.569	3.5 \pm 11.0	8.5
O_{1a} vs O_{1b} (intra-obs)	0.978	-2.8 \pm 7.1	6.2	0.981	-0.1 \pm 5.8	4.5	0.896	-2.3 \pm 5.7	0.9
SRF	0.755	-0.2 \pm 25.7	17.4	0.827	9.3 \pm 18.0	14.8	0.465	-11.5 \pm 15.4	12.8
BEASM-fully	0.704	13.4 \pm 30.6	22.9	0.713	18.0 \pm 25.8	22.5	0.731	-9.8 \pm 8.3	10.7
BEASM-semi	0.886	14.6 \pm 19.2	17.8	0.880	18.3 \pm 16.9	19.5	0.790	-9.4 \pm 7.2	10.0
U-Net 1	0.947	-8.3 \pm 12.6	10.9	0.955	-4.9 \pm 9.9	8.2	0.791	-0.5 \pm 7.7	5.6
U-Net 2	0.954	-6.9 \pm 11.8	9.8	0.964	-3.7 \pm 9.0	6.8	0.823	-1.0 \pm 7.1	5.3
ACNN	0.945	-6.7 \pm 12.9	10.8	0.947	-4.0 \pm 10.8	8.3	0.799	-0.8 \pm 7.5	5.7
SHG	0.943	6.4 \pm 12.8	10.5	0.938	-3.2 \pm 11.3	8.2	0.770	-1.4 \pm 7.8	5.7
U-Net ++	0.946	-11.4 \pm 12.9	13.2	0.952	-5.7 \pm 10.7	8.6	0.789	-1.8 \pm 7.7	5.6
p-values		< 0.05			\ll 0.05			\ll 0.05	

* *corr* : Pearson correlation coefficient; *mae* : mean absolute error.

5.3.3 U-Net behavior

From the results given in table 9 and 11, it appears that the U-Net method has the most effective architecture among the tested EDN models in terms of trade-off between the number of parameters and the achieved performance for the particular task of 2D echocardiographic image analysis. To better analyze the behavior of this model, we set up several additional experiments whose results are provided in Fig. 22 and 23. For all these experiments, even if the acquisitions were optimized to perform LV_{EF} measurements (meaning that part of the LA may or may not be fully visible depending on the acquisitions), we also investigated the capacity of U-Net to segment the LA in addition to the LV_{endo} and LV_{epi} . Moreover, since the two tested U-Nets produced overall close geometrical and clinical scores, we only used in this part the U-Net 1 model since it requires considerably less parameters to learn. Finally, all the given metrics were computed from both four and two-chamber views and at ED and ES time instants to facilitate the interpretation of the results.

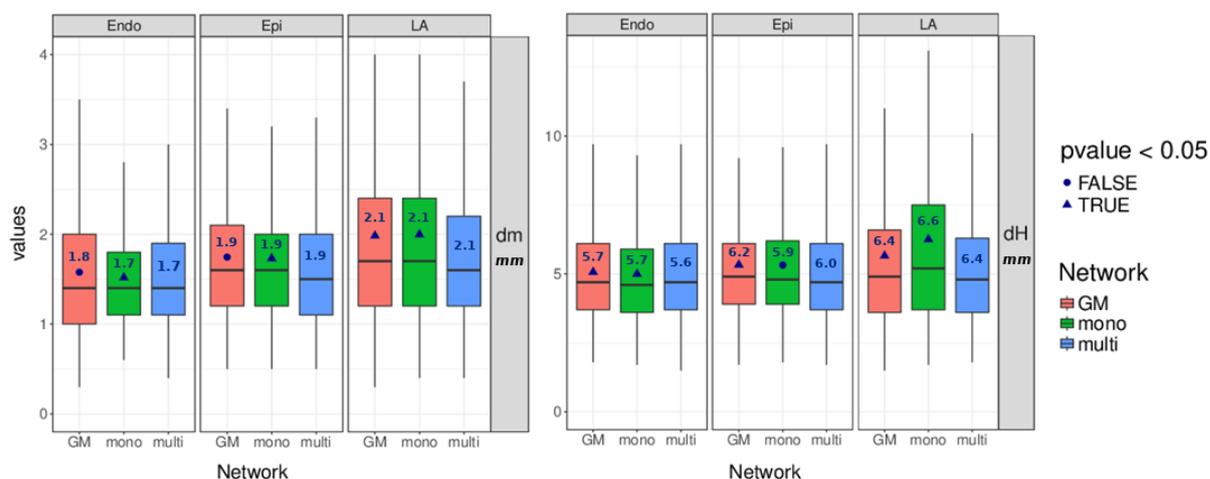


FIGURE 22 – Tukey box plots computed from the geometrical results of the U-Net 1 architecture for three different schemes (*GM* for learning to simultaneously segment all three structures from good and medium image quality, *mono* for learning to segment one structure from all image quality, *multi* for learning to simultaneously segment all three structures from all image quality). Blue numbers correspond to mean values computed from each set of measurements. p-values are based on the Wilcoxon signed rank test computed with the *multi* strategy as reference.

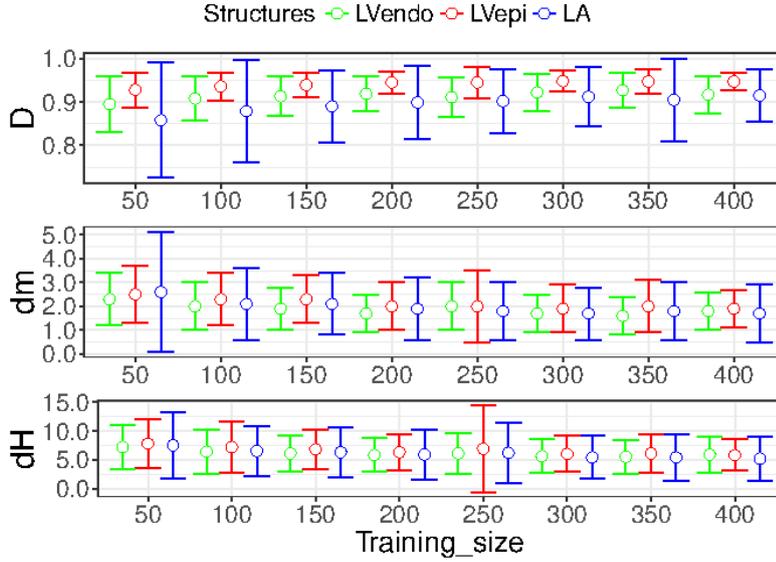


FIGURE 23 – Box plots results computed from the U-Net 1 architecture for three different schemes (GM for the learning from good & medium image quality, mono for the learning of one structure, multi for the learning of the three structures at the same time). Blue figures displayed under each box correspond to mean values computed from each set of measurements. p-values are based on the Wilcoxon signed rank computed with the multi strategy as reference.

Mono versus multi-structures approach : We assessed the influence of learning strategies on the quality of the segmentation of the LV_{Endo} , LV_{Epi} and LA. In particular, we trained 4 models with the same U-Net 1 architecture but with different training sets including all image quality, *i.e.* one network trained on predicting only the LV_{Endo} , one the LV_{Epi} , one the LA, and one all structures. Results on the full dataset are plotted in green and blue in Fig. 22 and are referred to as *mono* and *multi*. From the derived box plots, one can see that, unrelated to the structure, the mono and multi-structures approaches produced very close results even if the corresponding differences are statistically different. These results show that, with the proposed implementations, learning the segmentation of one structure (*e.g.* LV_{Endo}) in the context of the others (*e.g.* LV_{Epi} & LA) does not improve significantly the results compared to learning the segmentation of the structure alone. This hints at designing dedicated architectures and/or loss functions to better exploit the contextual information provided in the segmentation masks. Furthermore, even if the segmentation of the LA structure is challenging compared to LV_{Epi} and LV_{Endo} due to acquisition conditions, the U-Net 1 manages to get close results both in terms of mean absolute distance (mean d_m equals to 1.7, 1.9 and 2.1 mm for the LV_{Endo} , LV_{Epi} and LA respectively) and average Hausdorff distance (mean d_H equals to 5.6, 6.0 and 6.4 mm for the LV_{Endo} , LV_{Epi} and LA respectively).

The effect of poor quality images : We investigated in Fig. 22 the influence of involving images of poor quality during the training phase. Based on a multi-structures scheme, we trained two U-Net 1 models with the same architecture, one using the full training dataset not caring for image quality (plotted in blue and referred as *multi* in Fig. 22) and one using the training dataset restricted to patients having good and medium image quality (plotted in red and referred as *GM* in Fig. 22). From the obtained box plots, one can observe that the two different strategies produced very close results even if the corresponding differences are

mostly statistically significant (apart for the d_m metric for the LV_{Endo} and LV_{Epi}). These results suggest that the 19% (94 patients) of poor quality images *i)* do not bring additional information (supporting that the remaining deep learning issues are weakly linked to image quality); *ii)* do not decrease performance compared to a model trained on the 406 patients with good and medium image quality. This result suggests that poor image quality, in itself, does not complicate the segmentation task as much as could be expected and that encoder-decoder based techniques are able to cope with the variability in image quality found in echocardiography.

Influence of the size of the training database We studied in Fig. 23 the influence of the size of the training dataset on the quality of the segmentation of the LV_{Endo} , LV_{Epi} and LA structures. To this aim, we set up 8 different experiments, where the same fold 5 and 6 were respectively used as test and validation sets. As for the training set, starting from 50 patients, we added for each new experiment 50 additional patients until 400 patients was reached for the last trial. In each experiment, the same U-Net 1 architecture was used and optimized in the same way to derive the best performing parameters from the validation set. Moreover, the number of training epochs was proportionally lowered to ensure that each network went through the same number of iterations. From this figure, one can first observe an overall improvement of all metrics for the three cardiac structures with the increasing number of patients in the training set. Interestingly, while the improvement between 50 to 200 patients is quite pronounced (*e.g.* a decrease of d_H for the LV_{Endo} from 7.2 mm to 5.8 mm), one can observe a change in the evolution of the performance of the U-Net 1 method from 250 patients. Indeed, for this particular value, results worsen a bit, which may be explained by the bias brought by the validation and test data as we are not doing cross-validation in this experiment. Moreover, from this value, the d_m scores seem to stabilize around 1.8 mm for the LV_{Endo} , 1.9 mm for the LV_{Epi} and 1.8 mm for the LA structure. The same conclusions can be made for the Dice metric, with a convergence value around 0.920 for the LV_{Endo} , 0.947 for the LV_{Epi} and 0.909 for the LA structure. As for the d_H metric, while some improvement can still be observed from 250 to 400 patients for the LV_{Epi} and LA structures (1.1 mm and 1.0 mm for the LV_{Epi} and the LA, respectively), it is not obvious to draw the same conclusion for the LV_{Endo} structure since the decrease of its corresponding value is less pronounced (0.2 mm). In the light of these results, the U-Net 1 implementation performs better than the state-of-the-art non-deep learning methods after training with only 50 patients. Moreover, this method needs at least 250 patients during the training phase to reach highly competitive results, which can be slightly improved with a larger training set.

Influence of the expert annotations : We investigated in Fig. 24 the influence of the expert annotations during the training phase. To this aim, we trained three models on fold 5 from the same U-Net 1 architecture based each time on the manual contouring from a different annotator. The validation fold was kept the same for each experiment to avoid any bias error. The models were then evaluated on the remaining 400 patients annotated by cardiologist O_{1a} . From this figure, one can observe that the best scores for the three structures are obtained for the model trained on the annotations of cardiologist O_{1a} , who performed the manual contouring on the test and validation sets. This observation is consistent with the inter-variability results provided in table 9 and 11. It confirms that cardiologists have consistent differences in their way of contouring images and that an EDN has the capacity to learn a specific way of segmenting.

Runtime performance : The two U-Nets were implemented in Python with the same version of the TensorFlow and Keras libraries and an Nvidia Tesla M60 GPUs (8 Go RAM). Because of the larger number of trainable parameters involved in the U-Net 2 solution (see table 7 and 8), the running time of the two networks is different. For the training phase, the time

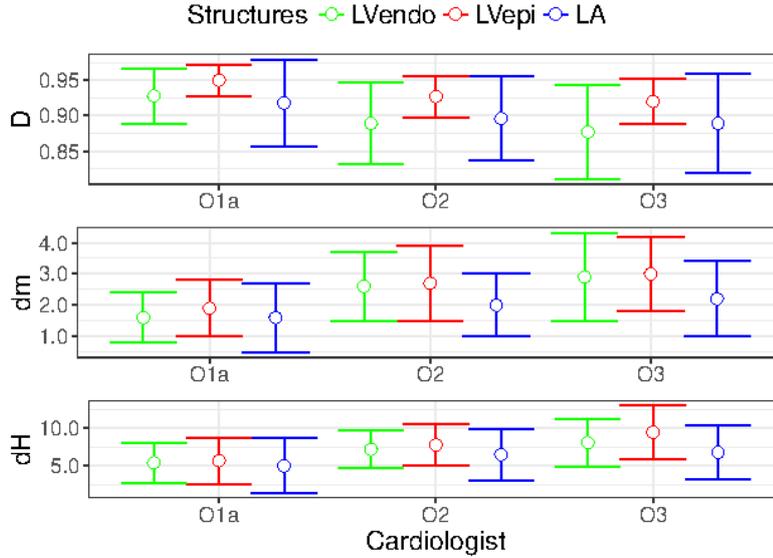


FIGURE 24 – Geometric scores of the three cardiologist-specific models on 400 patients (1600 images)

required to train on 400 patients is $24 \pm 5 \text{ min}$ and $73 \pm 1 \text{ min}$ for the U-Net 1 and U-Net 2, respectively. At test time, the segmentation of a single image takes $0.09 \pm 0.03 \text{ ms}$ and $0.14 \pm 0.06 \text{ ms}$ for the U-Net 1 and U-Net 2, respectively.

5.4 Discussion

5.4.1 Statistical differences between U-Net 1 and 2 results

From table 9 and 11 it has been observed that although U-Net 1 and 2 have very similar performances, their results were judged most of the time as being statistically different by the Wilcoxon Signed Rank Test ($p\text{-value} < 0.05$). This can be explained by the fact that when the number of samples is quite high, any slight but consistent deviation between the two distributions will make the difference statistically significant. In our study, since we work on a large scale dataset, most of the statistical tests were performed on a large amount of samples (for instance table 9 and 11 involve more than 800 paired observations for each statistical test), encouraging situations where the differences between results produced by two methods are recognized as statistically significant (even if the evaluated distributions are very close).

5.4.2 Inter and intra-observer variability

To further assess the quality of the EDN segmentation results, we added in table 9 the inter and intra-observer variability measurements computed from fold 5 (restricted to 40 patients with good and medium image quality). Concerning the inter-observer variability, the corresponding Dice scores vary between 0.82 and 0.93, the d_m between 2.2 mm and 4.0 mm and the d_H between 6.0 mm and 8.8 mm. The LV_{Epi} is the most difficult structure to annotate while both LV_{Endo} and LV_{Epi} are harder to contour at ES than at ED. One should also note the large d_m value of 4.0 mm between observer 1a and 3 for the LV_{Endo} structure at ES. This illustrates *i*) the difficulty in getting coherent manual annotations between experts from daily clinical practice

data; *ii*) the difficulty for the experts to use unfamiliar software for the analysis; *iii*) the needs to provide interactively the volumetric results to the experts for instant comparison (this was not done during the manual annotations); *iv*) the difficulty in contouring some data acquired with non-standard views. Concerning the intra-observer variability, one can observe that the results obtained on all the segmentation scores are better than the inter-observer ones, with a mean difference of 1.5 mm for the d_m metric and 2.6 mm for the Hausdorff distance. This illustrates the high consistency of manual contouring from experienced cardiologists, even on challenging data. Those results also provide important information on the limits to reach in order to consider that a machine learning algorithm faithfully reproduces the expertise of one cardiologist.

In table 11, we also reported the inter and intra-observer variability measurements computed from fold 5 (restricted to 40 patients with good and medium image quality) for the LV_{EDV} , LV_{ESV} and LV_{EF} metrics. From these results, one can observe that the experts reached good agreements for the estimation of the LV_{EDV} and LV_{ESV} with mean correlation scores of 0.92 and 0.91, respectively. However, the LV_{EF} results are worse with a mean correlation value of 0.67. This reveals the extreme difficulty in getting consistent fully manual annotations from ED to ES and between clinicians. It also illustrates the need for semi- or fully-automatic solutions to get higher temporal coherency, as illustrated by the higher LV_{EF} scores obtained by the semi-automatic BEASM method (0.79) and the EDN approaches (0.79 on average). Concerning the intra-observer variability, results are much more consistent with mean correlation scores of 0.98, 0.98 and 0.90 of the LV_{EDV} , LV_{ESV} and LV_{EF} metrics, respectively.

5.4.3 U-Net versus more sophisticated encoder-decoder architectures

Table 9 and 11 underlines that U-Net results are very close to those obtained by more sophisticated architectures. This is surprising as one might expect that more complex deep learning designs would improve results, at least marginally. As for ACNN, similar scores may be explained by the simple shapes encountered in 2D echocardiography. Indeed, the reference contours drawn by the experts involve truncated ellipse-like shapes whose information seems to be easily learned by the different EDNs. As a result, the anatomical constraint of the ACNN does not bring any additional value during the segmentation process, leading to similar or even slightly lower performance due to the regularization effect (which can lead to simpler shapes than expected). Concerning SHG and U-Net++, the similar scores may be explained by the results in Fig. 23. From this figure, we observed that U-Net reaches a plateau in terms of its performances when training on more than 250 patients. This suggests that the capacity of a U-Net is sufficient to generalize well on CAMUS dataset. Thus, the increase of the network capacity through the SHG or the U-Net++ architectures does not bring any improvement.

5.4.4 Accuracy of EDNs at delineating the LV_{Endo} , LV_{Epi} and LA structures

Segmentation results given in table 9 show that the five EDN implementations clearly outperform the state-of-the-art fully and semi-automatic non-deep-learning methods. In particular, while also learning from annotated data, SRF does not perform as consistently as the EDNs. Concerning the deformable model-based BEASM, even if it integrates the annotated information through a shape prior, this method produces overall significantly less accurate segmentation results. It thus appears from this study that a well-designed EDN can reach impressive segmentation scores in echocardiography. Interestingly, the EDN results are better than the inter-observer scores, on all structures and metrics. Although further investigations shall be made to validate this assertion, the obtained results tend to show that, when properly

trained, deep learning techniques are able to reproduce manual annotations with high fidelity. The results presented in this pilot study should thus stimulate the community to set up public multi-centric and multi-vendor datasets in echocardiography with annotations from cardiologists having passed high level consensus criteria. It is also interesting to note that EDN results are slightly worse than the intra-observer scores, on all structures and metrics (apart for the d_m metric for the LV_{Epi} at ED). This reveals that even if EDNs produce remarkable results, there still exists room for improvement to faithfully reproduce the manual annotations of one expert, taking into account its variability due to the ultrasound image quality.

In complement, we counted the number of cases for which the EDNs produced results outside the inter-observer variability, *i.e.* a d_m value higher than 3.5 mm and 4.0 at ED and ES, respectively and a d_H value higher than 8.2 mm and 8.8 at ED and ES, respectively. From this experiment, we found that 18% of the segmentations produced by both U-Nets, ACNN or SHG can be seen as outliers. This value goes up to 30 % for U-Net++. For comparison purpose, the outliers rate from two series of annotations on fold 5 produced by the same expert O_{Ia} is equal to 13%. Even if the overall performances of the EDNs are remarkable, this confirms the interest of still improving deep learning solution to produce highly reliable segmentation results on daily clinical practice data.

5.4.5 Accuracy of EDNs at estimating clinical indices

Clinical scores provided in table 11 show that EDNs produce results below the inter-observer scores for all the metrics. It thus appears that the evaluated EDNs are serious candidates to automatically produce trustworthy estimations of the LV_{EDV} and LV_{ESV} indices, on par with medical expertise. Concerning the LV_{EF} , even if the results are better than the inter-observer scores, a correlation value of 0.82 (for the best performing method) appears too low in comparison with the intra-observer value of 0.90 to consider the automatic estimate of this index as sufficiently robust to be dependable in clinical practice. The lower LV_{EF} scores compared to LV_{EDV} and LV_{ESV} measures can be partially explained by the lack of temporal coherency in the tested EDN implementations. Indeed, for each patient, the ED and ES frames are viewed as two independent images, potentially generating less efficient estimation of the corresponding LV_{EF} measures. Numerous deep learning strategies that integrate temporal coherence such as the recurrent neural networks (the Long Short Term Memory - LSTM - model being one of the famous network of this family) has been described. The integration of such concepts into the U-Net formalism seems to be a solution of interest in order to make the LV_{EF} estimation more accurate.

5.5 Conclusions

In this study, we introduced the largest publicly-available and fully-annotated dataset for 2D echocardiographic assessment (to our knowledge). The CAMUS dataset, containing 2D apical four-chamber and two-chamber view sequences acquired from 500 patients, is made available for download at <https://camus.creatis.insa-lyon.fr/challenge/>. A dedicated Girder⁸ on-line platform has been setup for new result submissions at <https://camus.creatis.insa-lyon.fr/challenge/>. Thanks to this dataset, the following new insights were underlined :

8. <https://girder.readthedocs.io>

- Encoder-decoder networks produced highly accurate segmentation results in 2D echocardiography ;
- Among the different tested architectures, U-Net appeared to be most effective in terms of trade-off between the number of parameters and the achieved performance ;
- The reasons for the lack of improvement of the more sophisticated networks (ACNN, SHG and U-Net++) compared to U-Net was addressed ;
- U-Net reached a plateau in terms of its performances when training on more than 250 patients but still continued to improve, implying that though 250 patients was enough to generalize well on CAMUS, it has the potential to integrate additional variability ;
- U-Net showed impressive robustness to variability, especially to image quality. Considering the wide range of image quality involved in echocardiography, this result is another positive element to consider encoder-decoder-based techniques as a solution of choice to solve the problem of 2D echocardiographic image segmentation ;
- U-Net learned to reproduce a specific way of contouring ;
- The segmentation and clinical results of the encoder-decoder networks were all below the inter-observer scores ;
- The segmentation and clinical results of the encoder-decoder networks were close to but slightly worse than the intra-observer scores. This reveals that even if encoder-decoder networks produced remarkable results, there is still room for improvement to faithfully reproduce the manual annotation of a given expert.

6 4th contribution : deep convolutional neural network study on cardiac MRI

In this section we present a study that we realized to measure how far the state-of-the-art deep learning method can go at assessing cardiac MRI, *i.e.* segmenting the myocardium and the two ventricles as well as classifying pathologies. In the wake of the 2017 MICCAI-ACDC challenge, we report results from deep learning methods provided by nine research groups for the segmentation task and four groups for the classification task from the ACDC database described in Sec. 3.3. This work has been the subject of one main publication [33].

6.1 Motivations

We proposed a new dataset called ACDC (see Section 3.3) which led to the organization of an international MICCAI challenge in 2017. The richness of the dataset as well as its tight bound to every-day clinical issues has the potential to enable machine learning methods to fully analyze cardiac MRI data. ACDC has a larger scope than previous cardiac datasets as it includes manual expert segmentation of the right ventricle (RV) and left ventricle (LV) cavities, and the myocardium (epicardial contour more specifically). ACDC also contains patients from five different medical groups namely : dilated cardiomyopathy (DCM), hypertrophic cardiomyopathy (HCM), myocardial infarction with altered left ventricular ejection fraction (MINF), abnormal right ventricle (ARV) and patients without cardiac disease (NOR). The overarching objective of this study is to provide answers to the following four questions :

1. How accurate recently proposed segmentation methods are at delineating the LV, RV and myocardium given clinical MR images ?
2. How accurate recently proposed classification methods are at predicting the pathology of a patient given clinical MR images ?
3. When methods fail, where do they fail ?
4. How far are we from "solving" the problem of automatic CMRI analysis ?

6.2 Evaluated methods

The ACDC dataset was made available during the "Automatic Cardiac Diagnosis Challenge (ACDC)" workshop held in conjunction with the 20th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), on September 10th, 2017 in Quebec City, Canada. After having publicly invited people to participate to this challenge, 106 accounts were created on the challenge website. Ten teams uploaded meaningful results within the allotted time for the segmentation contest, while 4 teams participated in the diagnosis contest.

6.2.1 Architecture for cardiac multi-structure segmentation

A summary of the ten architectures involved in this study is provided in table 12. Nine methods implemented a deep convolutional architecture, most of which a U-Net like networks [225] analyzing the 3D data slice by slice. The only exception is the method by Tziritas and Grinias [267] which implemented a Chan-Vese level-set method followed by a MRF graph cut segmentation method and spline fitting to smooth out the resulting boundaries.

TABLE 12 – Overview of methods evaluated during the ACDC challenge.

Reference *	Contest	Method	Remarks
Baumgartner <i>et al.</i> [28]	S	2D U-Net	Tested several architectures, the best one being a 2D U-Net with a cross-entropy loss
Isensee <i>et al.</i> [111]	S	2D+3D U-Net	Ensemble of 2D and 3D U-Net architectures with a Dice loss
Jang <i>et al.</i> [113]	S	2D M-Net	Use of a weighted cross-entropy loss function
Khened <i>et al.</i> [126]	S	Dense U-Net	2D U-Net with dense blocks and an inception first layer
Patravali <i>et al.</i> [194]	S	2D U-Net	Tested several architectures, the best one being a 2D U-Net with a Dice loss
Rohé <i>et al.</i> [223]	S	SVF-Net	Multi-atlas strategy where the registration module is realized using an encoder-decoder network
Tziritas and Grinias [267]	S	Levelset+MRF	Chan-Vese levelset followed by graph cut and a B-Spline fitting to smooth out results
Wolterink <i>et al.</i> [286]	S	Dilated CNN	Feed-forward CNN but with dilated convolution operations
Yang <i>et al.</i> [289]	S	3D U-Net	Use of 3D U-Net but with residual connections instead of the usual concatenation operator
Zotti <i>et al.</i> [304]	S	2D Grid-Net	Use of a Grid-Net architecture with an automatically-registered shape prior
Cetin <i>et al.</i> [43]	C	SVM	Use of physiological and radiomic (shape, intensity and texture) features
Isensee <i>et al.</i> [111]	C	RF	Extract a series of instant and dynamic features; use an ensemble of 50 multilayer perceptrons
Khened <i>et al.</i> [126]	C	RF	Extract 11 features from seg. results + patient height/weight; trained a 100-trees RF classifier
Wolterink <i>et al.</i> [286]	C	RF	Extract 14 features from seg. results + patient height/weight; trained a 1000-trees RF classifier

* S : Segmentation contest ; C : Classification contest ; SVM : Support Vector Machine ; RF : Random Forest ; MRF : Markov Random Field.

Four papers re-used the U-Net architecture. Baumgartner *et al.* [28] tested the U-Net and the FCN architectures with various hyper parameters. They also tested the impact of using 2D and 3D convolution layers as well as a training Dice loss versus a cross-entropy loss. Their best architecture ended up being a U-Net with 2D convolution layers trained with a cross-entropy loss. Isensee *et al.* [111] implemented an ensemble of 2D and 3D U-Net architectures (with residual connections along the upsampling layers). Concerning the 3D network, due to large inter slice gap on the input images, pooling and upscaling operations are carried out only in the short axis plane. Moreover, due to memory requirements, the 3D network involves a smaller number of feature maps. Both networks were trained with a Dice loss. Similar to Baumgartner’ study, Patravali *et al.* [194] tested a 2D and 3D U-Net trained with different Dice and cross entropy losses. From their experiments, the best performing architecture was a 2D U-Net with a Dice loss. Finally, Yang *et al.* [289] implemented a 3D U-Net but with residual connections instead of the usual concatenation operator. They also used pre-trained weights for the downsampling path using the C3D network known to work well on video classification tasks [264]. Their network was trained with a multi-class Dice loss.

Four papers used a modified version of the U-Net. Jang *et al.* [113] implemented a ”M-Net” [162] architecture whose main difference with U-Net resides in the feature maps of the decoding layers which are concatenated with those of the previous layer. The corresponding network was trained with a weighted cross-entropy loss. Khened *et al.* [126] implemented a dense U-Net. Their method starts by finding the region of interest with a Fourier transform followed by a Canny edge detector on the first harmonic image and compute an approximate radius and center of the LV with a circular Hough transform on the edge map previously generated. They then use a U-Net with dense blocks instead of basic convolution block to make the system lighter. The first layer of this network also corresponds to an inception layer. The system was trained with a sum of Dice and cross-entropy losses. Rohé *et al.* [223] developed a multi-atlas algorithm that first registers a target image with all images in the training dataset. The registered label fields are then merged with a soft fusion method using pixel-wise confidence measures. The registration module implements an encoder-decoder network called SVF-Net [224]. Finally, Zotti *et al.* [304] implemented a ”Grid Net” architecture which corresponds to a U-Net with convolutional layers along the skip connections. The architecture also registers a shape prior which is used as additional features map before performing the final decision. The model was

trained with a four term loss function.

Wolterink *et al.* [286] is the only team that implemented a CNN without an encoder-decoder architecture. Instead, they used a sequence of convolutional layers with increasing levels of kernel dilation to ensure that sufficient image context was used for each pixel's label prediction. This CNN was fed simultaneously with spatially corresponding ED and ES 2D slices while the output of the network was split in two, one softmax for ED and one for ES.

6.3 Solutions for automatic cardiac diagnosis

Three participants of the segmentation challenge used their segmentation result to extract features for cardiac diagnosis. Isensee *et al.* [111] extracted a series of instants and dynamic features from the segmentation maps and used an ensemble of 50 multilayer perceptrons (MLP) and a random forest to perform classification. Khened *et al.* [126] used 11 features, 9 derived from their segmentation map in addition to the patient weight and height. From those features, they trained a 100-trees random forest classifier. Wolterink *et al.* [286] extracted 14 features (12 from the segmentation maps + patient weight and height) and used a five-class random forest classifier with 1,000 decision trees.

Cetin *et al.* [43] were the only one to involve a semi-automatic segmentation method to manually extract the contours of the cardiac structures. Based on those contours, they computed 567 features including physiological features (*e.g.* height and weight) and radiomic features such as shape-based features, intensity statistics, and various texture features. To prevent their method from overfitting, they selected the most discriminative features and used SVM for classification.

6.4 Results

6.4.1 Evaluation metrics

In order to evaluate the tested methods in a fair and reproducible manner, we customized a dedicated Girder⁹ on-line platform¹⁰. This platform is now available and will be maintained and kept open as long as the data remains relevant for clinical research. Based to this platform, the performance of state-of-the-art methods are compared both from a geometrical and a clinical standpoint. This implies the use of a complementary set of metrics as described hereunder [134].

6.4.2 Geometrical metrics

In order to measure the accuracy of the segmentation output (LV endocardium, myocardium or RV endocardium) provided by a given method, the Dice metric and the Hausdorff distance were used. The corresponding formula can be found in Sec. 5.3.1. Concerning the Hausdorff distance, as opposed to several MRI cardiac segmentation papers which report 2D Hausdorff distances [16, 213, 253], we report the 3D d_H , which allows an intrinsic management of the missing segmentation problem on the end slices.

9. <https://girder.readthedocs.io/en/latest/>

10. <http://acdc.creatis.insa-lyon.fr/>

6.4.3 Clinical performance

We also implemented three indices for the clinical parameters, namely the correlation (*corr*), the bias and the standard deviation (*std*) values. These three metrics are computed from the measurements of : *i*) the ED volumes (LV_{EDV} and RV_{EDV} expressed in mL/m^2 for the LV and RV, respectively); *ii*) the ejection fractions (LV_{EF} and RV_{EF} expressed in percent for the LV and RV, respectively); *iii*) the myocardium mass (MY_{Mass} expressed in g/m^2 and calculated in diastole). The combination of the bias and standard deviation also provides useful information on the corresponding limit of agreement values. Let us mention that these geometrical and clinical metrics are complementary in the sense that a good score on one metric does not inevitably imply a good score on other metrics. This property is fundamentally important to prevent our system from unexpectedly favoring some methods over others. For instance, a low EF error does not always mean a good delineation of the ED and ES ventricle since EF relies on the difference between the ED and ES volumes. As such, a method that would systematically over- or underestimate the size of a ventricle in the same order at both ED and ES would potentially have a low EF bias, a low mean average error and a high EF correlation, but at the same time a low Dice score and a large Hausdorff distance.

6.5 Segmentation Challenge

Table 13 shows the segmentation testing accuracy (50 patients) for all 10 algorithms. The red values correspond to the best scores for each metric while the blue values correspond to the methods that are one pixel away from the top method. We use this color code to underline the closeness between the involved methods. This one-pixel criterion is a range of agreement of 2.3 mm for the Hausdorff distance (the maximum in-plane diagonal distance between two pixels : $\sqrt{(1.66^2) * 2}$) and 0.02 for the Dice metric (the average Dice score between the segmentation map of a method and the same segmentation map dilated or eroded by 1 pixel). This one pixel criterion comes from the fact that the two experts gave themselves a one pixel error margin such that two annotations were considered identical when their 2D Hausdorff distance was smaller or equal than one pixel. From these results, one can see that the 2D-3D U-Net ensemble model proposed by Isensee *et al.* [111] is overall the top performing method (the corresponding code is publicly available through the following link¹¹). This approach is closely followed by other methods which are less than one pixel away from it, especially for the LV and RV at ED. For instance, Baumgartner *et al.*, Jang *et al.*, Zotti *et al.*, and Khened *et al.* are within the range of agreement of the top performing method for 9 of the 12 metrics. As for the none deep-learning method by Tziritas and Grinias, it is relatively far away from the top, especially for the RV and the MYO. To allow a detailed analysis of the results, a set of segmentation output are provided in Fig. 25 to 29. This should help better assess the quality of the best approaches.

Table 15 contains the clinical metrics for all 10 methods. As for the segmentation part, red values correspond to the best scores for each metric. Blue values correspond to the methods with a p-value larger than 5% compared to the best method (we used an unequal variances two-sample t-test). For the clinical indices, Khened *et al.* [126] globally outperforms the other approaches with 14 metrics out of 20 close to the top performing method (*i.e.* red and blue metrics). In terms of correlation metrics, most of the methods obtained highly accurate results with values above 0.96 for the volumes. Methods also get good LV_{EF} results with high correlation scores, a bias close to zero (0.8% on average), a small mean absolute error (3.2% on average) and small

11. <https://github.com/MIC-DKFZ/ACDC2017>

TABLE 13 – Segmentation accuracy of the 10 evaluated methods on the testing dataset. **Red** is the best method, and **blue** are the methods within the range of agreement (Dice index of 0.02 and Hausdorff distance of 2.26 mm from the best).

Methods *	ED						ES					
	LV		RV		Myo		LV		RV		Myo	
	D	d _H										
	val.	mm										
Isensee <i>et al.</i> [111]	0.968	7.4	0.946	10.1	0.902	8.7	0.931	6.9	0.899	12.2	0.919	8.7
Baumgartner <i>et al.</i> [28]	0.963	6.5	0.932	12.7	0.892	8.7	0.911	9.2	0.883	14.7	0.901	10.6
Jang <i>et al.</i> [113]	0.959	7.7	0.929	12.9	0.875	9.9	0.921	7.1	0.885	11.8	0.895	8.9
Zotti <i>et al.</i> [304]	0.957	6.6	0.941	10.3	0.884	8.7	0.905	8.7	0.882	14.1	0.896	9.3
Khened <i>et al.</i> [126]	0.964	8.1	0.935	14.0	0.889	9.8	0.917	9.0	0.879	13.9	0.898	12.6
Wolterink <i>et al.</i> [286]	0.961	7.5	0.928	11.9	0.875	11.1	0.918	9.6	0.872	13.4	0.894	10.7
Jain <i>et al.</i> [194]	0.955	8.2	0.911	13.5	0.882	9.8	0.885	10.9	0.819	18.7	0.897	11.3
Rohé <i>et al.</i> [223]	0.957	7.5	0.916	14.1	0.867	11.5	0.900	10.8	0.845	15.9	0.869	13.0
Tziritas-Grinias [267]	0.948	8.9	0.863	21.0	0.794	12.6	0.865	11.6	0.743	25.7	0.801	14.8
Yang <i>et al.</i> [289]	0.864	47.9	0.789	30.3	N/A	N/A	0.775	53.1	0.770	31.1	N/A	N/A

* ED :End diastole ; ES : End systole ; LV : Endocardial contour of the left ventricle ; RV : Endocardial contour of the right ventricle ; Myo : Epicardial contour of the left ventricle (myocardium) ; D : Dice Index ; d_H : Hausdorff distance.

standard deviations (4.3%). The most difficult clinical metric to estimate is the EF of the RV with a correlation score of 0.9 for the best method.

A joint analysis of table 13 and 15 reveals that results on the myocardium (especially at ES) are those that vary the most. This may be partially explained by the fact that an accurate myocardium segmentation implies the precise delineation of two walls instead of one for the LV and RV. Methods also struggle with the RV. The RV often has the highest Hausdorff distances, the lowest Dice scores, the lowest correlation values, and the largest biases. To further underline this observation, we recorded in table 14 the percentage of patients for which the predicted EF is less than 5% away from the ground-truth (5% is often considered as an acceptable error margin [39]). While the top six methods accurately predict the LV ejection fraction for $\approx 87\%$ of the patients, that number drastically goes down to $\approx 59\%$ for the RV.

TABLE 14 – Percentage of patients with an EF error lower than 5%.

Methods	LV	RV
Isensee <i>et al.</i> [111]	92 %	68%
Jang <i>et al.</i> [113]	88 %	60%
Rohe <i>et al.</i> [223]	88 %	34%
Zotti <i>et al.</i> [304]	84 %	60%
Khened <i>et al.</i> [126]	84 %	56%
Baumgartner <i>et al.</i> [28]	84 %	54%
Wolterink <i>et al.</i> [286]	80 %	38%
Jain <i>et al.</i> [194]	68 %	54%
Tziritas-Grinias [267]	66 %	38%
Yang <i>et al.</i> [289]	58 %	32%

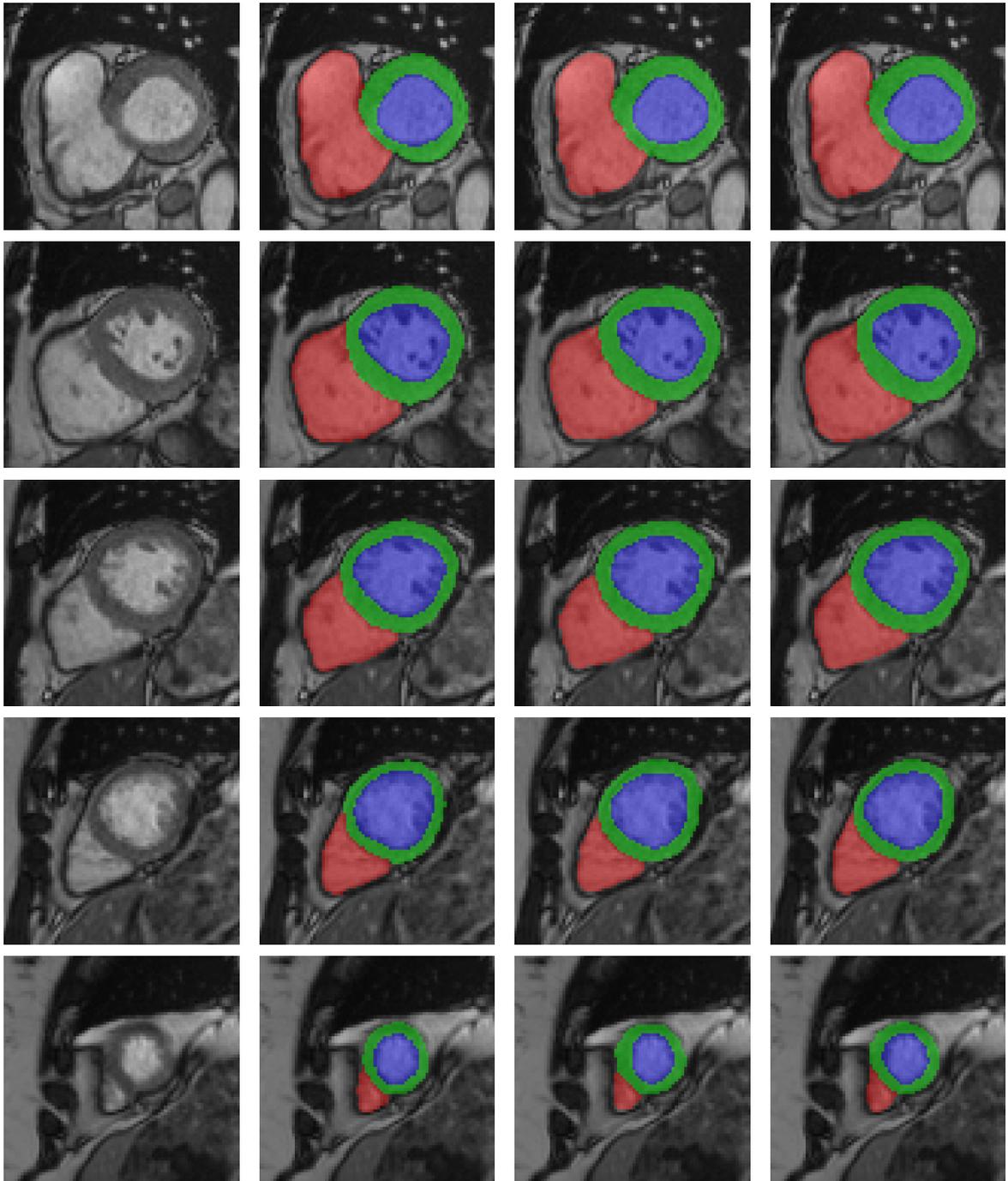


FIGURE 25 – Automatic segmentation of a healthy subject (NOR) from the testing database. Each row corresponds to a given slice of the same volume at end diastole (ED). From top to bottom : basal slice toward apical slice. From left to right : input image, ground-truth, Isensee *et al.* (winner of the segmentation contest), Khened *et al.* (winner of the classification contest). For this patient (full volume at ED), Isensee *et al.* obtained a Dice score of 0.97, 0.92 and 0.96 and a Hausdorff distance of 2.2 mm, 3.1 mm, and 3.1 mm, respectively for the left ventricle, myocardium and right ventricle. Khened *et al.* got a Dice score of 0.97, 0.92 and 0.97 and a Hausdorff distance of 3.1 mm, 3.1 mm, and 3.1 mm, respectively for the left ventricle, myocardium and right ventricle.

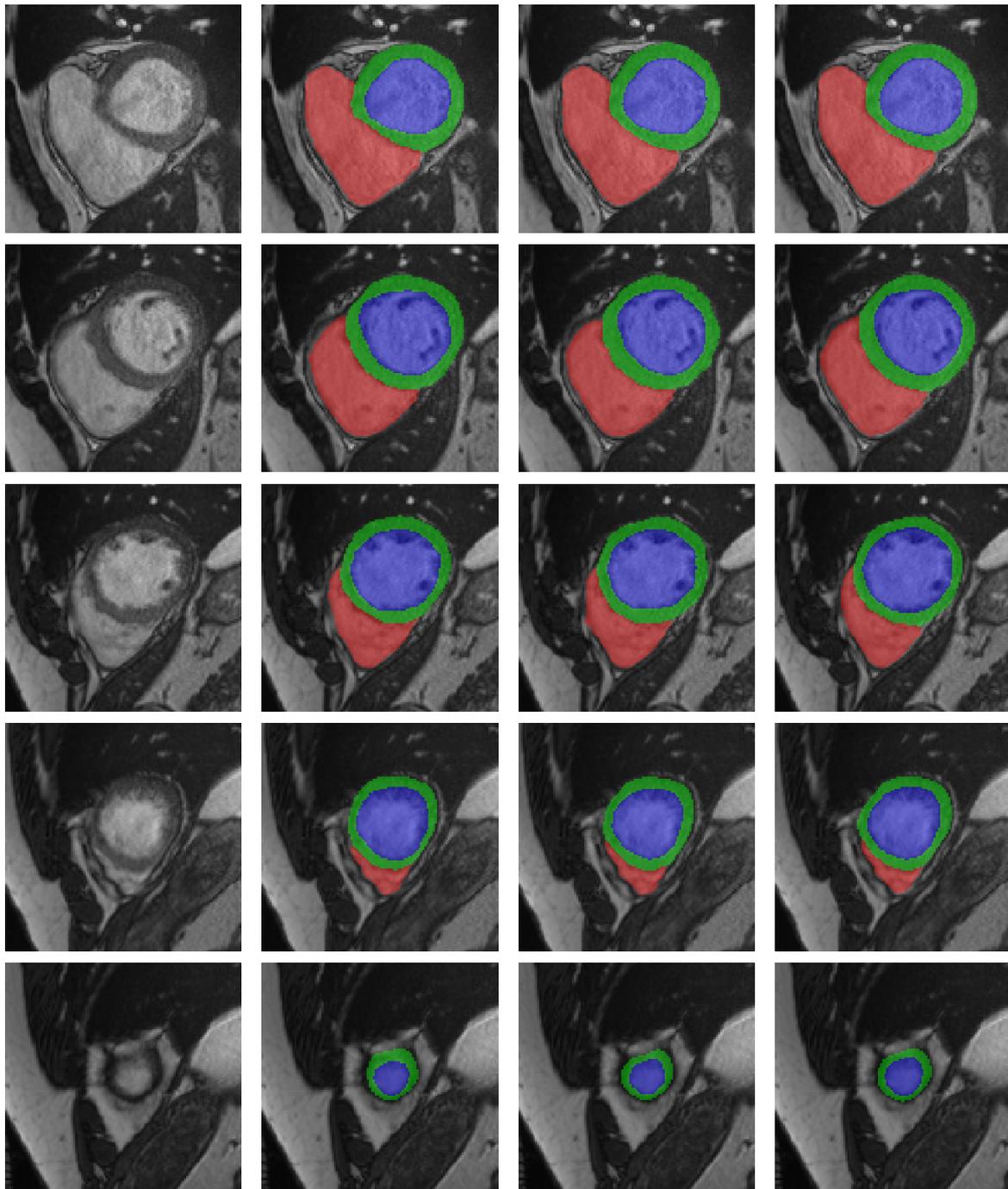


FIGURE 26 – Automatic segmentation of a patient with a myocardial infarction (MINF) with altered left ventricular ejection fraction from the testing database. Each row corresponds to a given slice of the same volume at end diastole. From top to bottom : basal slice toward apical slice. From left to right : input image, ground-truth, Isensee *et al.* (winner of the segmentation contest), Khened *et al.* (winner of the classification contest). For this patient (full volume at ED), Isensee *et al.* obtained a Dice score of 0.97, 0.91 and 0.97 and a Hausdorff distance of 2.7 mm, 3.9 mm, and 4.9 mm, respectively for the left ventricle, myocardium and right ventricle. Khened *et al.* got a Dice score of 0.97, 0.90 and 0.95 and a Hausdorff distance of 3.1 mm, 4.1 mm, and 10.7 mm, respectively for the left ventricle, myocardium and right ventricle.

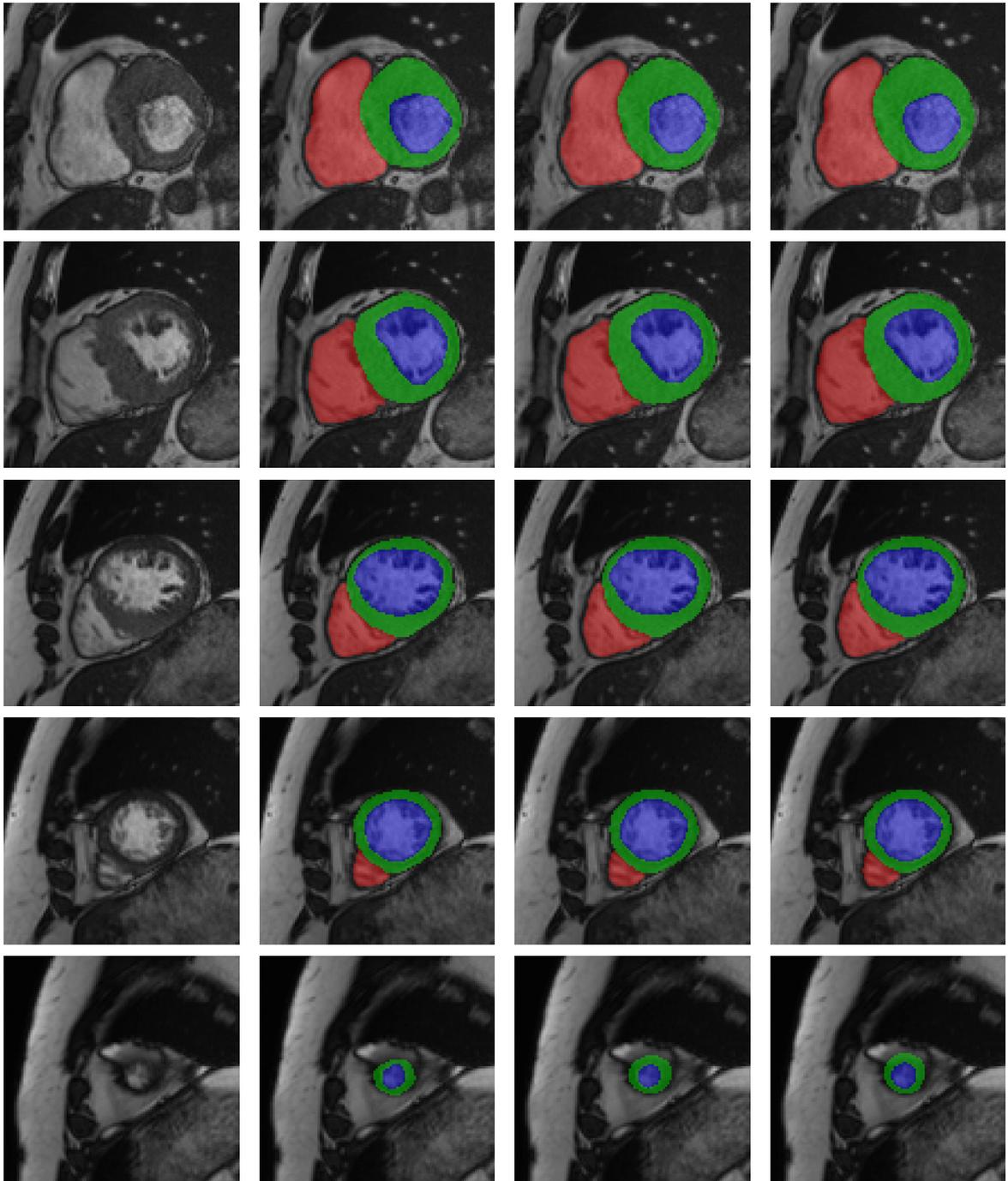


FIGURE 27 – Automatic segmentation of a patient with an hypertrophic cardiomyopathy (HCM) from the testing database. Each row corresponds to a given slice of the same volume at end diastole. From top to bottom : basal slice toward apical slice. From left to right : input image, ground-truth, Isensee *et al.* (winner of the segmentation contest), Khened *et al.* (winner of the classification contest). For this patient (full volume at ED), Isensee *et al.* obtained a Dice score of 0.92, 0.89 and 0.97 and a Hausdorff distance of 11.6 mm, 10.5 mm, and 5.9 mm, respectively for the left ventricle, myocardium and right ventricle. Khened *et al.* got a Dice score of 0.97, 0.92 and 0.96 and a Hausdorff distance of 4.6 mm, 10.0 mm, and 11.7 mm, respectively for the left ventricle, myocardium and right ventricle.

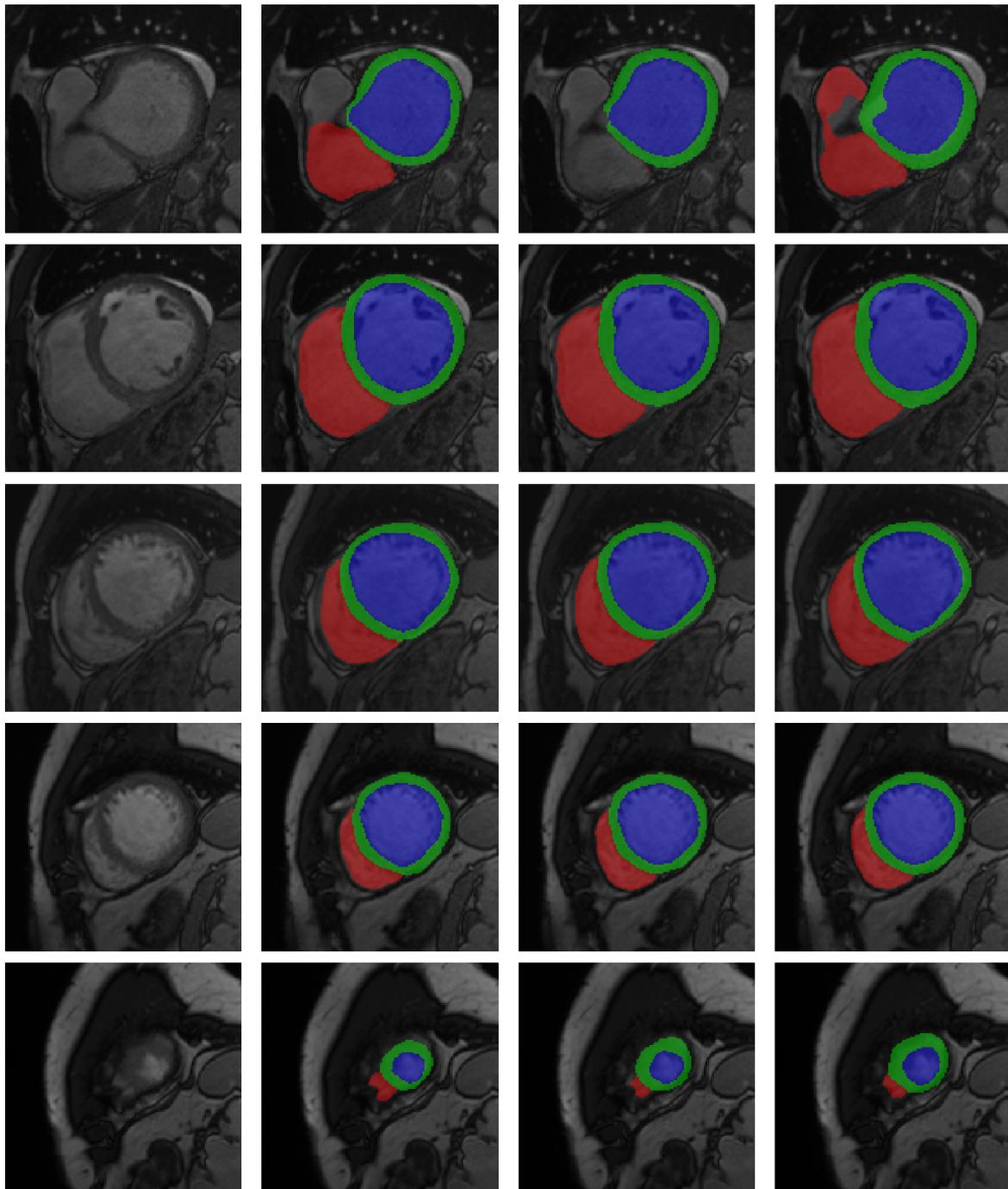


FIGURE 28 – Automatic segmentation of a patient with dilated cardiomyopathy (DCM) from the testing database. Each row corresponds to a given slice of the same volume at end diastole. From top to bottom : basal slice toward apical slice. From left to right : input image, ground-truth, Isensee *et al.* (winner of the segmentation contest), Khened *et al.* (winner of the classification contest). For this patient (full volume at ED), Isensee *et al.* obtained a Dice score of 0.98, 0.91 and 0.90 and a Hausdorff distance of 4.1 mm, 4.3 mm, and 10.7 mm, respectively for the left ventricle, myocardium and right ventricle. Khened *et al.* got a Dice score of 0.97, 0.87 and 0.92 and a Hausdorff distance of 10.4 mm, 11.6 mm, and 10.5 mm, respectively for the left ventricle, myocardium and right ventricle.

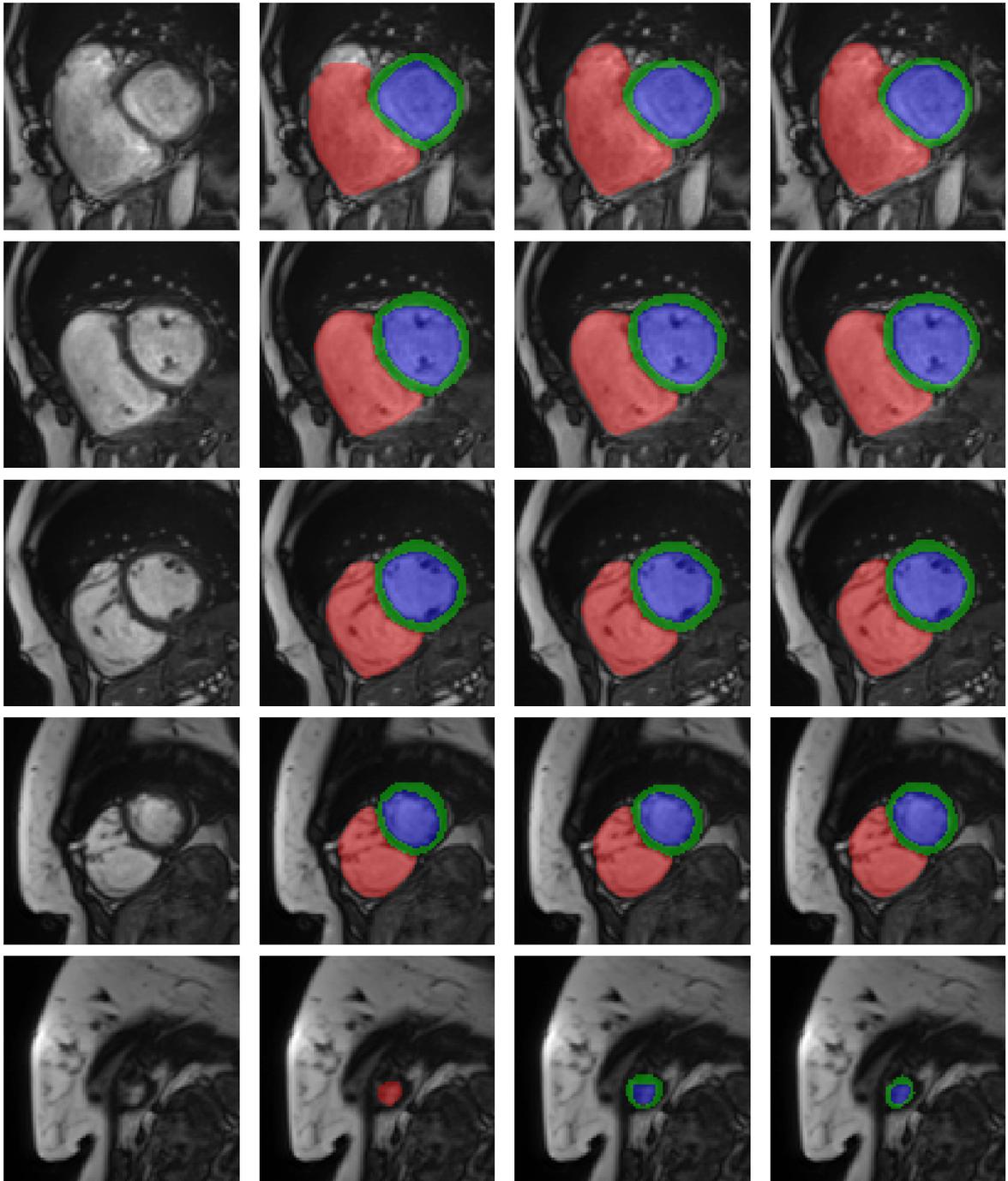


FIGURE 29 – Automatic segmentation of a patient with abnormal right ventricle (ARV) from the testing database. Each row corresponds to a given slice of the same volume at end diastole. From top to bottom : basal slice toward apical slice. From left to right : input image, ground-truth, Isensee *et al.* (winner of the segmentation contest), Khened *et al.* (winner of the classification contest). For this patient (full volume at ED), Isensee *et al.* obtained a Dice score of 0.97, 0.89 and 0.97 and a Hausdorff distance of 24.3 mm, 24.1 mm, and 12.8 mm, respectively for the left ventricle, myocardium and right ventricle. Khened *et al.* got a Dice score of 0.97, 0.86 and 0.95 and a Hausdorff distance of 24.3 mm, 24.1 mm, and 12.9 mm, respectively for the left ventricle, myocardium and right ventricle.

TABLE 15 – Clinical metrics for the 10 evaluated methods on the testing dataset. **Red** is the best method, and **blue** are the methods within a p-value larger than 0.05 according to bias and std measurements.

Methods *	LV _{EDV}			LV _{EF}			RV _{EDV}			RV _{EF}			MY _{Mass}		
	corr	bias±σ	mae	corr	bias±σ	mae	corr	bias±σ	mae	corr	bias±σ	mae	corr	bias±σ	mae
	val.	ml.	ml.	val.	%	%	val.	ml.	%	val.	%	%	val.	g.	g.
Khened <i>et al.</i> [126]	0.997	0.6 ± 5.5	4.2	0.989	-0.5 ± 3.4	2.5	0.982	-2.9 ± 12.6	8.4	0.858	-2.2 ± 6.9	5.3	0.990	-2.9 ± 7.5	6.3
Isensee <i>et al.</i> [111]	0.997	2.7±5.7	5.1	0.991	0.2 ± 3.1	2.1	0.988	4.4±10.8	7.9	0.901	-2.7 ± 6.2	4.7	0.989	-4.8 ± 7.6	7.3
Zotti <i>et al.</i> [304]	0.997	9.6±6.4	10.3	0.987	-1.2 ± 3.6	2.7	0.991	-3.7 ± 9.2	7.4	0.872	-2.2 ± 6.8	5.4	0.984	-12.4±9.0	13.1
Jain <i>et al.</i> [194]	0.997	9.9±6.7	10.8	0.971	1.7±5.5	4.1	0.945	5.6±22.2	15.0	0.791	6.8±8.1	8.3	0.989	11.6±8.1	11.9
Wolterink <i>et al.</i> [286]	0.993	3.0±8.7	6.8	0.988	-0.5 ± 3.4	2.5	0.980	3.6 ± 15.2	10.9	0.852	-4.6±6.9	6.6	0.963	-1.0 ± 14.6	10.0
Jang <i>et al.</i> [113]	0.993	-0.4 ± 8.7	6.0	0.989	-0.3 ± 3.3	2.3	0.986	-10.8±11.6	12.1	0.793	-3.2 ± 8.3	6.3	0.968	11.5±12.9	14.1
Baumgartner <i>et al.</i> [28]	0.995	1.4 ± 7.6	6.1	0.988	0.6 ± 3.4	2.6	0.977	-2.3 ± 15.1	11.1	0.851	1.2 ± 7.3	5.7	0.982	-6.9±9.8	9.8
Rohé <i>et al.</i> [223]	0.993	4.2±8.6	7.5	0.989	-0.1 ± 3.2	2.6	0.983	7.3±13.4	11.7	0.781	-0.7 ± 9.9	7.8	0.967	-3.4 ± 13.3	10.3
Tziritas-Grinias [267]	0.992	2.0 ± 11.7	8.5	0.975	-1.6 ± 5.0	4.3	0.930	18.6±25.4	24.8	0.758	-0.5 ± 9.1	7.1	0.942	-28.9±28.0	30.3
Yang <i>et al.</i> [289]	0.894	12.2±32.0	27.5	0.926	1.5 ± 8.7	6.1	0.789	47.3±41.9	48.7	0.576	8.8±23.2	15.7	N/A	N/A	N/A

* LV_{EDV} : End diastolic left ventricular volume ; LV_{EF} : Left ventricular ejection fraction ; RV_{EDV} : End diastolic right ventricular volume ; RV_{EF} : Right ventricular ejection fraction ; MY_{Mass} : Myocardial mass in diastole ; mae : mean absolute error

Table 16 presents an overview of the classification performance of the 4 evaluated methods. Due to the small number of samples (50 patients), the scores have to be considered with care since a miss-classification causes an accuracy drop of 2%. From this table, one can see that Khened *et al.* [126] obtained nearly perfect results with 48 patients correctly classified. The confusion matrix of this approach is shown in Fig. 30. Please note that the best approach is closely followed by the next two methods which obtained an accuracy of 92%.

TABLE 16 – Results on the classification challenge.

Methods		Accuracy
Authors	Architectures	
Khened <i>et al.</i> [126]	Random Forest	0.96
Cetin <i>et al.</i> [43]	SVM	0.92
Isensee <i>et al.</i> [111]	Random Forest	0.92
Wolterink <i>et al.</i> [286]	Random Forest	0.86

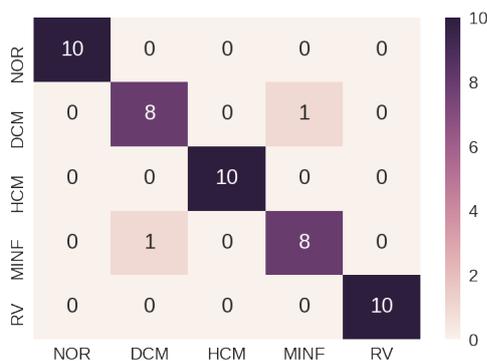


FIGURE 30 – Confusion matrix of the winner of the classification challenge [126].

Let us mention that although MINF and DCM are visually similar, MINF implies a local lack of myocardial contraction as opposed to DCM. Moreover, for DCM, the LV must exceed 100 mL/m². This is why machine learning methods have been able to successfully differentiate these pathologies.

6.6 Discussion

How far are we from solving the CMRI analysis problem ? Automatic classification results (healthy subjects and patients with 4 different pathologies), showed that the best methods are very close to each other with an accuracy above 92%. Although these observations have to be validated on more patients, it appears from this study that well designed machine learning techniques can reach near perfect classification scores.

However, conclusions are not so straightforward for the segmentation task. While results obtained on the LV are competitive, it appears that the same level of accuracy is still difficult to obtain for the RV and the MYO. It is thus important to assess the performance of the top methods relatively to the experts variability. Unfortunately, the actual version of the ACDC

TABLE 17 – Dice and Hausdorff distances for *i*) inter- and intra- observers *ii*) the average of every submitted deep learning (DL) methods and *iii*) the winner of the segmentation challenge. **RED** corresponds to results within or above the inter-observer variation. The last 5 lines correspond to metrics computed without the apical and the basal slices.

Methods *	ED						ES					
	LV		RV		MYO		LV		RV		MYO	
	D	d _H										
	val.	mm										
O_{1a} vs O_2 (inter-obs)	0.956	5.6	0.930	12.6	0.870	6.7	0.898	8.1	0.866	14.0	0.891	7.6
O_2 vs O_{1b} (inter-obs)	0.950	6.2	0.931	12.1	0.868	7.2	0.895	8.5	0.861	14.1	0.886	8.0
O_{1a} vs O_{1b} (intra-obs)	0.967	4.0	0.957	7.6	0.900	5.1	0.941	5.4	0.930	9.1	0.917	6.0
Average DL methods vs GT	0.965	7.6	0.947	13.2	0.906	10.1	0.927	9.2	0.886	15.2	0.898	10.9
Isensee <i>et al.</i> [111] vs GT	0.968	7.4	0.946	10.1	0.902	8.7	0.931	6.9	0.906	12.1	0.919	8.7
O_{1a} vs O_2 (inter-obs)	0.956	4.4	0.938	7.7	0.867	5.0	0.913	5.5	0.890	8.7	0.894	5.5
O_2 vs O_{1b} (inter-obs)	0.953	4.9	0.937	8.6	0.864	5.5	0.905	5.8	0.898	9.4	0.886	6.1
O_{1a} vs O_{1b} (intra-obs)	0.971	3.1	0.960	5.8	0.905	3.6	0.950	3.9	0.940	6.9	0.923	4.4
Average DL methods vs GT	0.972	3.7	0.951	8.1	0.896	5.2	0.929	4.2	0.899	9.9	0.915	6.1
Isensee <i>et al.</i> [111] vs GT	0.972	3.7	0.969	6.4	0.910	4.6	0.945	4.2	0.912	8.6	0.930	5.1

* ED : End diastole ; ES : End systole ; LV : Endocardial contour of the left ventricle ; RV : Endocardial contour of the right Ventricle ; Myo : Myocardium contours ; D : Dice Index ; d_H : Hausdorff distance ; GT : Ground-truth.

dataset comes with one expert annotation per subject and does not provide any inter- or intra-observer error margin.

In order to evaluate the inter- and intra-observer variabilities, we asked the two experts O_1 and O_2 that jointly annotated the ACDC ground-truths to independently relabel the images of the 50 test subjects. O_1 annotated twice the images (we call those annotations O_{1a} and O_{1b}) one month apart while O_2 annotated the images once. The average geometric distance between O_{1a} , O_{1b} and O_2 are given in the first three lines of table 17. As one can see, the Dice scores oscillates between 0.86 and 0.96 and the HD between 4 mm and 14.1 mm. Without much surprise, the RV at ES is the most difficult region to annotate, even for experimented observers. It is also interesting to note that the Dice variations (especially for the inter-observer) are very close to that reported in a recent publication by Wenjia *et al.* [82]. As for the d_H values, the ones reported in table 17 are larger than those in Wenjia *et al.*'s paper due to the fact that our implementation of d_H accounts for the 3D structures of the heart. With an inter-slice thickness of 10 mm (in average), any slight lateral shift between two annotations greatly increases the d_H score.

Below the inter- and intra-observer results given in table 17, we provide *i*) the average geometrical metrics obtained by the deep learning methods involved in the challenge and *ii*) the scores obtained by Isensee *et al.*, the winner of the segmentation challenge. Interestingly, their Dice scores are all between the inter-observer and intra-observer scores. This suggests that state-of-the-art deep learning techniques have reached a plateau in the light of this metric. Although further investigations shall be made to validate this assertion (especially for images acquired from a set of more heterogeneous settings), the obtained results tend to show that, when properly trained, deep learning techniques are able to improve the Dice scores all the way to those of an expert. As for the d_H scores, methods are slightly above the inter-observer scores, but by only 2 to 3 mm.

TABLE 18 – Inter- and intra-observer variation of the mean absolute error of the LV_{EDV} , RV_{EDV} and MY_{Mass} . Below, the winner of the challenge and the average deep learning methods compared with the ACDC ground-truth. **RED** are result between the inter- and intra-observer variance.

	LV_{EDV}	RV_{EDV}	MY_{Mass}
	ml.	ml.	g.
O_{1a} vs O_2 (inter-obs)	10.4	9.2	12.6
O_2 vs O_{1b} (inter-obs)	10.8	9.5	11.5
O_{1a} vs O_{1b} (intra-obs)	4.6	5.7	6.2
Average methods vs GT	7.1	10.6	10.4
Isensee et al. vs GT	5.1	7.9	7.3

In table 18, we put the inter- and intra-observer mean absolute errors computed from the LV_{EDV} , RV_{EDV} and MY_{Mass} metrics. From the given numbers, one can see that the inter- and intra-observer scores are very close to that reported by Wenjia *et al.* [82]. Moreover, the results obtained by Isensee *et al.* and the average deep learning methods are between the inter- and intra-observer scores.

Where do methods fail ? In the light of the results reported so far, it appears that top deep learning segmentation methods are in the range of human expects according to the Dice scores and the clinical metrics but still 2 to 3 mm away from experts in regards of the 3D Hausdorff distance. One may thus wonder where do methods fail ? One hypothesis can be that hearts suffering from a pathology may be more difficult to segment. To verify this assumption, we broke down in Fig. 31 the average Dice and Hausdorff metrics for each pathology obtained by the challengers on the test set (we remind that each pathological case corresponds the same amount of patients, both for the training and the testing phases). As one can see, there is no pathology for which methods systematically fail. For instance, while the HCM Dice score is somewhat low for the LV-ES (certainly due to the difficulty to see the cardiac cavity), it is larger than the other pathologies for MYO-ES and MYO-ED. Also, contrary to what one might think, images from healthy subjects (NOR) are not easier to segment than those from pathological cases as the scores relative to this group get the largest Hausdorff distances for the LV-ED and LV-ES.

Another hypothesis would be that 1.5T images are more difficult to segment than 3T CMR images due to an intrinsic lower SNR. However, after careful analysis of segmentation results, we found no particular differences between 1.5T and 3T results, as illustrated in table 19. One reason for this could be explained by the fact that both 1.5T and 3T images were included in the training set thus allowing neural networks to learn a representation specific to both magnetic fields.

Another hypothesis commonly accepted in the community is that slices next to the valves and/or the apex of the ventricle are more difficult to segment due to partial volume effect with surrounding structures. To investigate this assumption, we computed the total number of 2D

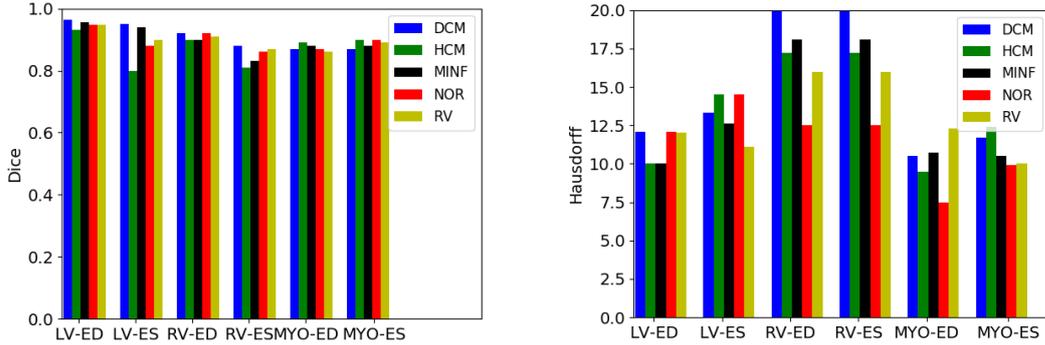


FIGURE 31 – Average Dice index and Hausdorff distances for every method reported in Table 13 broken down for every pathology.

TABLE 19 – Dice scores of the winner of the segmentation challenge [111] on the 1.5T and 3T CMR images taken from the testset.

	ED			ES		
	LV	RV	MYO	LV	RV	MYO
1.5T	0.97	0.95	0.90	0.93	0.90	0.92
3T	0.97	0.94	0.91	0.94	0.88	0.92

segmentation results produced by each method for which the LV, MYO or RV had a Dice score below 0.70. The corresponding results are summarized through the histogram in Fig. 32, where the x-axis stands for the slice position (from the valves on the left to the apex on the right). Please note that since the number of slices varies from one patient to another, we stacked the 2D segmentation result of each method and made a 3D volume. Each volume was then resized to 10 slices with a nearest neighborhood interpolation method. From this figure, one can see that segmentation results obtained next to the valves and the apex are far more error prone. In particular, we notice almost 50% of results with very low Dice score at the apex (often because LV/MYO/RV are very small at that position). As for the base, we observe that methods often struggle to differentiate between the RV, the LV, the atria and the surrounding structures (cf. Fig. 33). We also put in table 17 the Dice and Hausdorff metrics computed without the apical and basal slices. While the Dice scores are almost identical with and without the end slices, the Hausdorff distance decreases significantly, sometimes by a factor of two for the learning methods. Interestingly, the learning methods fall within the inter- and intra-observer variabilities (apart for Hausdorff metric for the RV at ES) which shows that segmenting apical and basal slices is far more difficult, even for experts.

Finally, it is worth pointing that the use of a larger database than the one involved in this project might help in resolving the listed remaining issues. For instance, the UK Biobank [199] may be a serious candidate for this purpose. We thus see the UK Biobank and our database as complementary with the strong potential to offer materials for upcoming research studies.

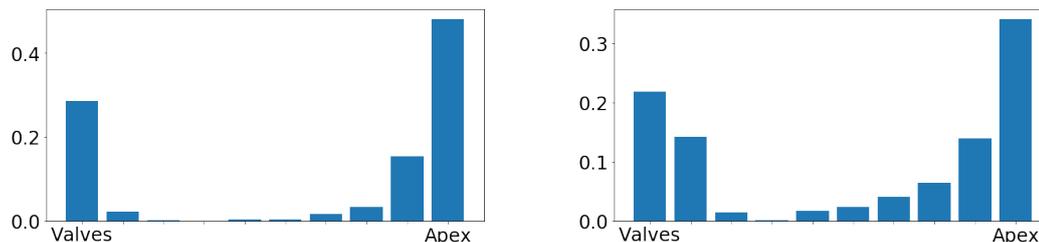


FIGURE 32 – Histogram of degenerated slices ED (left), and ES (right).

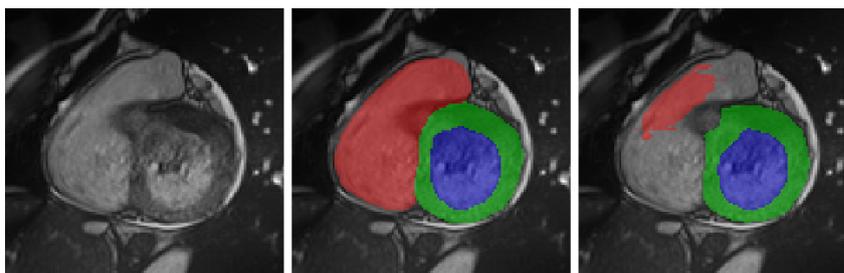


FIGURE 33 – Typical degenerated result at the base of the heart. [Left] input image ; [Middle] ground truth ; [Right] prediction.

For the need of a new metric : Results reported so far suggest that top deep learning methods are very close to the inter-observer variability. However, the visual inspection of their segmentation results reveal that unlike experts, deep learning methods sometimes generate anatomically impossible results as shown in Fig. 33. Interestingly, the metrics used to gauge performances seem resilient to such abnormalities. In order to measure the number of anatomically impossible results, one of our expert visually screened the test results by Isensee *et al.* [111]. This revealed that results for 41 patients out of 50 had at least one slice with an anatomically impossible segmentation such as the RV disconnected from the MYO or the LV cavity in contact with the background (several detailed examples are given in the supplementary materials available in the supplementary files /multimedia tab). Those 41 patients had problematic results for 1.6 slices on average, most of them located next to the valves or the apex. This clearly underlines the fact that clinical and geometrical metrics used to assess results have important limits and that methods within the inter-observer variability may still be error-prone. This suggests the need for new evaluation metrics before one may claim that methods have reached the accuracy of an expert.

6.7 Clinical implications

Results presented thus far suggest that we are at the eve of cracking the nut of fully automatic CMRI analysis. This would allow to reduce the time spent on analyzing raw data so conclusions of the examination could be provided to the patient before leaving the radiology department. In today's clinical practices, the latest systems provide pre-filled radiologic reports with an integrated automatic speech recognition technology so doctors can dictate the various physiological and technical parameters. An automatic CMRI analysis software could thus easily be integrated within this framework. That being said, further investigations are still required before such software gets approved by accreditation agencies (CE mark, FDA, ISO, etc.) and

get integrated in MRI consoles. Also, although classification software get near-perfect results, the use of a "diagnostic black box" could not be integrated as-is in a clinical practice. Along with the pathology prediction, a medical report must always contain the physiological reasons for which the patient was diagnosed in a certain way. This calls for cardiac parameters such as EF, volumes, and mass estimated by a segmentation method which, in the context of deep learning approaches, may sometimes fail at the apex and the base and even produce anatomically impossible results. One shall also perform further analysis on images acquired by a wider variety of MRI scanners with different acquisition protocols to better assess the true generalization accuracy of machine learning algorithms.

Further research is also required on patient data suffering from other pathologies. Although we believe that some other pathologies such as inflammatory cardiomyopathy could be successfully diagnosed with the proposed machine learning methods, other (yet more complex) diseases such as congenital heart diseases or heart defect, would need dedicated studies.

6.8 Conclusions

ECG-gated sequences such as Cine-MRI allow for accurate analysis of left and right ventricular functions. The delineation of ventricular endocardium and epicardium allows the calculation of different parameters, such as LV_{EF} , RV_{EF} , myocardial mass, myocardial thickness, tele-systolic and tele-diastolic ventricular volumes. These measurements are an integral part of the exam interpretation by the radiologist and are necessary for the diagnosis of many cardiomyopathies. In this study, we have shown that state-of-the-art machine learning methods can successfully classify patient data and get highly accurate segmentation results. Results also reveal that the best convolutional neural networks get accurate correlation scores on clinical metrics and low bias and standard deviation on the LV_{EDV} and LV_{EF} , two of the most commonly-used physiological measures. However, methods are still failing at the base and the apex, especially when considering the Hausdorff distance.

7 Summary - Conclusions

From Sec. 3 to 6 we presented our strategy to tackle the problem of fully automatic and fast techniques for LV volumetric assessment through segmentation tasks in US and MR imaging.

To this purpose, we first proposed to put in place a set of benchmarked datasets to allow the community with a fair and reproducible comparison of state-of-the-art methods in segmenting cardiac images and extracting relevant clinical indices (*i.e.* the LV_{EDV} , LV_{ESV} , LV_{EF} , RV_{EDV} , RV_{EF} and MY_{mass}). In particular, we set up the following open-access datasets with the corresponding online evaluation platform :

- the CAMUS dataset containing 2D echocardiographic sequences (both two and four-chambers views) of 500 patients that were acquired during clinical exams in the same medical center with the same equipment. This dataset has been set up to evaluate how far state-of-the-art deep convolutional networks and others non-deep learning methods can go at segmenting cardiac structures (LV_{Endo} , LV_{Epi} and LA) as well as estimating clinical indices (LV_{EDV} , LV_{ESV} and LV_{EF}) ;
- the CETUS dataset containing 3D echocardiographic sequences of 45 patients that were acquired during clinical exams in three different centers with three different equipments (GE, Siemens and Philips). This dataset has been set up to evaluate how far state-of-the-art methods can go at segmenting LV_{Endo} structure as well as estimating clinical indices (LV_{EDV} , LV_{ESV} and LV_{EF}) ;
- the ACDC dataset containing cine MR sequences (multi-slice 2D cine MRI) of 150 patients that were acquired during clinical exams in the same medical center with two MRI scanners of different magnetic strengths (1.5T and 3.0T). The target population is composed of patients evenly divided into 5 classes with well-defined characteristics according to physiological parameters. This dataset has been set up to evaluate how far state-of-the-art deep convolutional network methods can go at segmenting cardiac structures (LV, MYO, RV), estimating clinical indices (LV_{EDV} , LV_{EF} , RV_{EDV} , RV_{EF} and MY_{mass}) as well as classifying the examination into 5 classes (normal case, heart failure with infarction, dilated cardiomyopathy, hypertrophic cardiomyopathy, abnormal right ventricle).

Thanks to those dataset, the following innovations and new insights were underlined :

1. *CAMUS dataset*

- Encoder-decoder networks produced highly accurate segmentation results in 2D echocardiography ;
- Among the different tested architectures, U-Net appeared to be most effective in terms of trade-off between the number of parameters and the achieved performance ;
- The reasons for the lack of improvement of the more sophisticated networks (ACNN, SHG and U-Net++) compared to U-Net was addressed ;
- U-Net reached a plateau in terms of its performances when training on more than 250 patients but still continued to improve, implying that though 250 patients was enough to generalize well on CAMUS, it has the potential to integrate additional variability ;
- U-Net showed impressive robustness to variability, especially to image quality. Considering the wide range of image quality involved in echocardiography, this result is

another positive element to consider encoder-decoder-based techniques as a solution of choice to solve the problem of 2D echocardiographic image segmentation ;

- U-Net learned to reproduce a specific way of contouring ;
- The segmentation and clinical results of the encoder-decoder networks were all below the inter-observer scores ;
- The segmentation and clinical results of the encoder-decoder networks were close to but slightly worse than the intra-observer scores. This reveals that even if encoder-decoder networks produced remarkable results, there is still room for improvement to faithfully reproduce the manual annotation of a given expert.

2. *CETUS dataset*

- The non-deep learning method that we developed in section 4 (named as BEASM) produced the current best scores for the segmentation of the LV_{Endo} in 3D echocardiography ;
- The greatest errors produced by the BEASM method occur on the apical region and on the anterior side of the LV ;
- A deep learning method with anatomical shape regularization [183] gets results that are quite close to those obtained with our method. The access to only 15 patients during the training phase illustrates the strong potential of deep learning techniques to analyze 3D echocardiographic images ;
- the best performing methods get accurate correlation scores on clinical metrics with low bias and standard deviation on the LV_{EDV} , LV_{ESV} and LV_{EF} , the LV_{EF} remaining the most difficult indices to estimate ;
- The segmentation and clinical results of the best performing methods are all slightly worse than the inter-observer scores. This reveals that there is still a need for improvement of the developed methods for the segmentation of the LV_{Endo} and the estimation of the corresponding clinical indices from 3D echocardiography.

3. *ACDC dataset*

- Deep convolutional networks produced highly accurate segmentation results in cine MRI ;
- Deep convolutional networks are still failing at the base and the apex, especially when considering the Hausdorff distance ;
- The best convolutional neural networks get accurate correlation scores on clinical metrics and low bias and standard deviation on the LV_{EDV} and LV_{EF} , two of the most commonly-used physiological measures ;
- Deep convolutional networks have still difficulties in estimating the RV_{EF} compared to the others tested indices ;
- The clinical results of the best convolutional neural network were all below the inter-observer scores ;
- The segmentation and clinical results of the best convolutional neural network were close to but slightly worse (apart for the Dice metric) than the intra-observer scores. This reveals that even if deep convolutional neural networks produced remarkable

results, there is still room for improvement to faithfully reproduce the manual annotation of a given expert ;

- Although these observations have to be validated on more patients, it appears from this study that well designed machine learning techniques can reach near perfect classification scores.

From these observations, it clearly appears that deep learning methods can be considered as a solution of choice to solve the problem of cardiac image segmentation and clinical indices estimation, both in 2D and 3D echocardiography and in cine MR imaging. Indeed, when the size of the dataset is enough (*i.e.* higher than 100 patients in the case of the proposed studies), both the segmentation and clinical results derived from well designed deep learning techniques were all below the inter-observer scores. However, in all the realized studies, the best performing deep learning networks failed to produce results in the range of the intra-variability scores. This should stimulate the community in still improving deep learning formalism in medical imaging to faithfully reproduce the manual annotation of a given expert without any anatomical failure.

My research on cardiac image segmentation was carried out through 4 PhDs funded by the French government (1 ministerial and 1 LABEX PRIMES grants) and Europe (2 PhDs financed from the ERC grant of professor Jan D’hooge). Results have led to 17 articles in international journals and 31 articles in international conferences. I also co-organized two challenges on this topic during MICCAI conferences (2014 and 2017) and setup three dedicated Girder on-line platforms which are still open for open-access dataset download and new result submissions (<https://www.creatis.insa-lyon.fr/EvaluationPlatform/CETUS/>, <https://acdc.creatis.insa-lyon.fr/> and <http://camus.creatis.insa-lyon.fr/challenge/>)

Cardiac motion estimation : toward robust strain estimation

1 Introduction

Cardiovascular diseases (CVD) are the number one cause of death and are expected to have an increasing prevalence in the coming years [227]. An estimated 17.5 million people died from CVDs in 2012, representing 31% of all global deaths [163]. Of these deaths, an estimated 7.4 million were due to coronary heart disease and 6.7 million were due to stroke. Assuming similar growth rates, by 2030 an estimated 23.6 million people will die from cardiovascular diseases in the world [279].

Heart disease, such as myocardial ischemia and ventricular dyssynchrony, may be identified and localized by the analysis of the cardiac motion and deformation. Early efforts used surgical implantation and tracking of radiopaque markers with X-ray imaging for quantifying the ventricular motion in canine hearts [91]. Such techniques are invasive and affect the regional motion pattern of the heart during the marker tracking process, thus not feasible clinically. With the advent of more advanced imaging techniques, non-invasive diagnosis techniques based on inspecting medical images were developed and have been widely used in clinical practice. Several modalities, such as echocardiography [123], magnetic resonance (MR) imaging and cardiac computed tomography (CT) have been applied to cardiac function analysis [166, 186, 255]. From medical images acquired by these modalities, physicians can monitor and assess the progression of CVD so that effective procedures for the disease treatment can be developed accordingly.

So far, cardiac motion and strain imaging have been primarily based on echocardiography and MR imaging. Cardiac CT has proven to be useful for obtaining a static segmentation at a high spatial resolution. It is also used for the segmentation of the coronary tree after contrast injection. Additionally, cardiac multi-detector computed tomography (MDCT) [226] is able to acquire dynamic images of the heart as a detailed source of information, it presents however several inherent limitations, if compared with echocardiography and MR imaging : 1) CT relies on ionizing radiation for delivering image contrast, which is harmful to human tissues ; 2) the temporal resolution of MDCT is lower, hence hindering cardiac motion tracking since fewer frames are available within the same cardiac cycle ; and 3) although the cardiac contour is prominent in CT images, the myocardial wall lacks a discernible pattern or "markers" that increases the difficulty for motion tracking [255]. As a result, cardiac CT is not as widely available as echocardiography and MR imaging for motion analysis. In our studies, we therefore concentrate on cardiac motion tracking methods from echocardiography and cardiac MR.

Traditionally, cardiac deformation is assessed from medical images by visual inspections

of an expert and each myocardial segment is then assigned a wall motion score. However, manual motion analysis similar to this is time-consuming and suffers from inter- and intra-observer variabilities. Automatic motion tracking softwares can help reducing the workload of the clinicians. Moreover, the medical conditions can be interpreted more objectively. Most of the motion tracking algorithms assess cardiac function through two categories of indexes : global and regional indexes. Global indexes include cavity volume, ejection fraction, cardiac output etc. They are important indexes that reveal the overall performance of the heart in terms of supplying oxygenated blood to the body. Regional indexes such as strain and left ventricle torsion may prove beneficial for the diagnosis of disease and evaluation of severity and the quantification of treatment effect [279]. Combining the global and regional analyses could lead to a reliable evaluation of the cardiac function.

Regarding the local motion tracking, echocardiography is one of the most common modality in clinics. The ultrasound is safe since no radiation is involved, real-time, portable and much cheaper compared to other imaging devices such as MR and CT. The speckle pattern inherent to ultrasound imaging can be used as natural markers that move with the myocardium. Speckle remains locally stable during a few consecutive frames. Therefore, the underlying myocardial motion can be accurately recovered by tracking them. Both the global ejection fraction and local myocardial strains can be further derived from the tracking field. Recently, with the development of 2D matrix transducer technology, 3D echocardiography is becoming more and more common in clinical facilities [5, 170]. Despite the reduced frame rate compared to 2D ultrasound, it still allows for an accurate quantification of cardiac motion [54]. Cardiac cine MR [232] is considered as a more advanced modality for assessing global ventricular functions in clinical routine. It has advantages over other modalities such as echocardiography and tagged MR [17] in that it provides clearer contrast between the myocardium and the blood pools. This facilitates the segmentation and the tracking of the borders of myocardium, *i.e.* the endocardium and the epicardium. From the trackings, global indexes such as left ventricle ejection fraction can be accurately calculated which allows an elementary assessment of cardiac function. It is more reproducible than echocardiography since it is easier to segment the myocardial borders. However, another characteristic of cardiac cine MR is that the images show nearly constant intensities within the myocardium, hampering somehow the quantification of local indexes such as motion and strain. Although reliable dense tracking field can still be computed by spatial regularizations, the absence of any local "markers" could reduce the tracking accuracy. Aside from ultrasound, cardiac tagged MR [17] is another imaging technique able to quantify reliable local myocardial deformations. In this modality, a specific tagging preparation pulse is composited with conventional MR imaging in order to deliberately create markers (tags) by saturating the tissue magnetizations. Those tags will move alongside the myocardium over the whole sequence. Deformation fields can therefore be accurately reconstructed from tracking those virtual markers. Tagged MR is considered as the gold standard for estimating regional myocardial deformations since it provides the most detailed information about local motion. However, its use in clinical practice is held back by the long acquisition time and the limited spatial resolution for distinguishing tissues.

Despite all those developments of imaging techniques, the assessment of cardiac function generally relies on global measures, such as left ventricle (LV) ejection fraction (EF) and global longitudinal strain (GLS) [243]. GLS quantified from speckle tracking echocardiography [38] is currently the best evaluated strain parameter [49][241]. Compared with global strains, regional strains are less reliable and reproducible as reported in [241][129][242]. Consequently, regional strain analysis is currently applied as a supplementary diagnosis method, but is still not ready

for full clinical implementation in spite of their potential impact in many applications[243][195]. One of the reasons is the fact that a well-established validation of regional strain imaging is still missing at the moment, despite the existence of a large number of dedicated quantification algorithms both in US [7] and tagged MR [299]. A comprehensive and thorough quantitative validation of these methods based on dense strain measurements would thus represent a major progress. Several groups already made some initial attempts towards this objective. Synthetic US [5, 94, 205], cine MR [205][94] and tagged MR [50][300] images were generated. They all adopted computer-generated synthetic sequences for validating motion tracking for one specific image modality. The interest in using synthetic images has been well detailed in several studies [5][205]. Their main advantage is the straightforward possibility of assessing the accuracy of algorithms by comparing their performance against the controlled (and thus known) ground-truth deformations. Nonetheless, the utility of such synthetic images depends heavily on the realism of the simulations. Lots of efforts have been made this last decade to improve the realism of cardiac synthetic images in each modality, as described in Sec. 3.

The remainder of the chapter is organized as follows.

- Sec. 2 : Review of state-of-the-art methods in cardiac motion estimation ;
- Sec. 3 : Review of state-of-the-art methods in cardiac simulation ;
- Sec. 4 : Overview of a novel pipeline specifically designed to simulate realistic synthetic US and MRI sequences from the same virtual patients ;
- Sec. 5 : Detailed description of an anatomical deformable model to efficiently quantify myocardial motion field and strain at once ;
- Sec. 6 : Conclusions of this chapter.

2 State-of-the-art in cardiac motion estimation

2.1 Cine MR

Cardiac cine MR is widely used in clinical practice due to its high soft-tissue contrast, especially between the myocardium and the neighboring blood pools. The endocardium and epicardium can therefore be accurately segmented. The temporal resolution is also high enough to do motion tracking over the cardiac cycle. By tracking the endo and epicardial borders over the whole cycle, global indexes such as left ventricle ejection fraction is easily quantified. It is widely accepted as the reference modality for quantifying global cardiac function. However, inside the muscle, cine MR shows nearly spatially constant intensities. Different from cardiac ultrasound, there are no natural markers that "tags" the myocardium. Dedicated strategies have been thus developed in order to quantify regional cardiac deformation despite the lack of natural tagging. They can be roughly classified into two categories.

Feature tracking

Hor *et al.* [106, 107] developed a feature tracking method applied to cardiac cine MR images. Endocardial and epicardial borders were manually drawn at ED frame. The myocardial border, as well as the columns of pixels radiating out from the endocardium, were then automatically propagated through the cardiac cycle by matching individual patterns that represent anatomical structures [107]. These were identified by the method of maximum likelihood between the regions of interest of consecutive frames. Myocardial strains were then computed from the tracking field. Similarly, Maret *et al.* tracked the myocardial borders from cine MR images and showed that displacement and strain in the radial and longitudinal direction could be used for the

detection of transmural scar [161]. Recently in [168], myocardial strain quantified from cine MR feature tracking was compared with tagged MR quantification results. The authors claimed that sufficient agreement was found between the two imaging modalities and feature tracking was considered as potentially feasible and rapid alternative. Similarly in [256] and [40], cine MR was shown capable of obtaining normal values for myocardial strain measurements on healthy volunteers and detecting myocardial fibrosis in pediatric hypertrophic cardiomyopathy.

Several other groups also tracked myocardial motion from cine MR images in a similar way. In [213], myocardial contours were first segmented by a B-spline explicit active surfaces technique, and then tracked throughout the heart cycle by an anatomically constrained optical flow method. Lin *et al.* proposed a generalized robust point matching framework to track the motion of left ventricle [176]. The displacements of myocardial borders obtained by robust point matching were further regularized by the free-form deformation model. Shi combined image-derived information (the tracked myocardial surfaces) and mechanical modeling of the myocardium and solved for deformations by finite-element models *et al.* [240].

Registration-based methods

Others regarded the motion tracking as a non-rigid registration problem. They aimed to estimate appropriate non-linear transformation fields, from which cardiac motion and strain were further derived. Sundar *et al.* used a 4D image registration method to estimate cardiac motion [252]. An attribute vector for every point in the image was used to contain information about intensity, boundary, and geometric moment invariants. Perperidis *et al.* proposed two B-spline-based free-form deformation methods for the spatiotemporal alignment of cardiac cine MR image sequences [198]. The two B-splines methods use the same transformation but differ in the order of the optimization process. One performed a combined optimization with spatial and temporal components. The other optimized each component separately. Cardiac deformations were then computed from the transformation field. Recently, Tsadok *et al.* applied non-rigid registration to cine MR images and observed that the regional longitudinal strain estimates correlated highly with tagged MR and STE [266].

2.2 Tagged MR

Tagging of myocardial tissue using MRI was first introduced by Zerhouni *et al.* [293]. The magnetization of the tissue is labeled or tagged by applying a spatially modulated saturation radiofrequency (RF) pulse before the actual MR imaging procedure. The tagging preparation sequence is played out at end-diastole upon detection of the ECG R-wave in planes perpendicular to the imaging plane. Since magnetization represents a characteristic tissue property, saturated regions are moving with the underlying tissue and displayed as black lines in the acquired images. By tracking these black lines, cardiac tissue can be tracked. Due to the T1 relaxation, the initially modulated magnetization tend to gradually return to its thermal equilibrium state. In the final image, this induces a progressive loss of contrast in the tags, known as tag fading [187]. As a result, the tags are only visible during a limited time period. Nonetheless, at 1.5T, the T1 relaxation time of healthy myocardium is about 850ms and therefore long enough to study cardiac contraction during the entire duration of a cardiac cycle. Since its introduction in the late 1980s, cardiac tagged MR has triggered the development of a wide spectrum of strain quantification methodologies. Some of them were specifically tailored for tagged MR. Some others are generic methodologies that were readily applied to tagged MR images. In the sequel, we focus on the methods that were specifically developed to tagged MR. Those methods can be

divided into four categories. These categories are rather separated, opening up the possibility of combining their respective advantages. Also, most of these methods were only applied to 2D tagged MR and adaptations to 3D are currently in an early stage [229], or focus on multiplanar [149, 189] rather than the volumetric acquisition protocol of Rutz *et al.* [229].

Detection-based methods

A first option retained by Young *et al.* and Amini *et al.* was to first detect the tags planes or intersections and then to track these extracted features. Amini *et al.* detected tag planes by B-spline surfaces and tracked sparse myocardial beads [10]. Young *et al.* used an active contour model to track the tag lines in images and then interpolated a 3D dense motion by a finite-element model [292]. In [291], the tagging stripes were detected by a Gaussian-shaped filter and matched by a line searching algorithm. The obtained sparse displacements were further regularized by a finite-element model defined in the local radial, longitudinal and circumferential directions. Chen *et al.* applied Gabor filters to detect tag line intersections which are further used to initialize a meshless deformable model [46].

HARP

HARP is another tagged MR-specific method that became a reference in this modality [186]. A first step in HARP is to compute phase images. They are computed by applying a band-pass filter in the Fourier domain. Back to the spatial domain, the tracking is then performed on phase (rather than intensity) images. The main rationale for substituting intensity by phase is to improve robustness against tag fading. The overall tracking procedure can be interpreted as an optical flow performed jointly on two “channels”, each channel corresponding to one tagging direction. Because the number of tagging directions equals the number of components of the displacement field, the aperture problem simply disappears. The estimation of HARP displacements was further improved when introducing the CSPAMM protocol, enabling the use of a wider k-space HARP filter [133]. However, unlike traditional optical flow algorithms, HARP does not integrate any spatial regularization. As strain is computed from the spatial derivatives of trajectories, one expects that outliers in the reconstructed motion field will yield high errors in the strain field. The HARP community also proposed several strategies for addressing the limited capture range of tagged images. Indeed, the periodicity of tagged images prevents to recover displacements exceeding one-half of the tag spacing. Osman proposed an initialization scheme by sorting the list of points to be tracked [186]. Points with expected smaller motion are tracked first and the tracking result is used as initialization to neighboring points. This concept was further developed in [150] where the order is optimized through solving for the shortest path on a graph where each edge has a cost related to phase continuity between neighboring voxels in the image. However, the integration of an order to the point-wise tracking is somehow dangerous. Indeed, “tag jump” errors will be propagated to all subsequent points in the list.

SinMod

The SinMod algorithm has recently emerged as an alternative to HARP [14]. In the local environment of each pixel, the intensity distribution is modeled as a sinusoid wave orthogonal to the tagging direction. Then both the local frequency and the inter-frame phase shift are computed. The ratio of these two quantities gives an estimate of the displacement when combining the different tagging directions. SinMod was shown to be as fast as HARP and have advantage in accuracy and robustness to image noise. Yet, the ability of tracking large myocardial motion (superior than half tag spacing) remains to be further verified. Also, the absence of any spatial regularization makes strain estimates very sensitive to tracking errors when extracting spatial

derivatives.

Gabor-based methods

Qian *et al.* proposed to compute strain directly from the image without tracking the myocardial motion [211]. The idea is to filter the intensity images through a Gabor filter bank for retrieving the local orientation and spacing of the tags. This gives a direct access to tissue deformation, without the need of computing displacements. This concept was further extended by Kause *et al.* where the deformation gradient was directly evaluated from the frequency covector fields [125]. The authors claim that it makes their strain estimate independent to any tracking result. However, they overlook that a tracking is always required for reporting strain evolution at all time points per material point.

2.3 Echocardiography

In the fields of cardiology and medical imaging, speckle tracking echocardiography (STE) is an echocardiographic imaging technique that analyzes the motion of tissues in the heart by using the naturally occurring speckle pattern in the myocardium when imaged by ultrasound insonification. 2D STE was introduced for estimating myocardial deformations [181, 219]. STE represents a noninvasive method of evaluating regional cardiac deformations, which is a valuable tool for the characterization and quantification of local myocardial deformations. The speckle pattern is a mixture of interference patterns and natural acoustic reflections, refraction and scattering [219]. The pattern is generated randomly and each region of the myocardium has a unique speckle pattern that allows the region to be traced from one frame to the next. The speckles stay relatively stable, at least during several frames when the acquisition frame rate is high enough compared to the dynamic of the structure of interest. The movement of the kernel across the image can thus be tracked. The speckles are therefore used as natural acoustic markers for "tagging" the myocardial motion during the cardiac cycle. By adopting a frame-to-frame tracking approach, the complete LV wall motion throughout the sequence can be accurately recovered. In principle, speckle can be tracked in both the axial and lateral dimensions. However, the axial resolution of the ultrasound is far better than the lateral. The tracking ability is thus decreased in the lateral direction. Moreover, the lateral resolution decreases with imaging depth, hence further limiting the tracking ability. Over the last decade, several speckle tracking techniques has been developed in order to quantify regional cardiac deformations. They can be roughly classified into four categories.

Block matching

B-mode block matching : Block matching algorithms such as [56, 69] represent the conventional implementations on many existing commercial systems. In these approaches, myocardial tissue is divided in blocks that are tracked throughout the cardiac cycle. For each local block, an exhaustive search is performed over a predefined search window in order to find the most similar patch at the next frame. Similarity can be defined in several ways. For instance, in [69], the authors chose the normalized cross correlation. The independently tracked patch-wise displacements are then regularized after block matching by interpolation and filtering, aiming to ensure a smooth estimate of myocardial motion.

RF block matching : The same concept was equally applied to the raw radiofrequency (RF) data (modulated signal derived at the first step of the beamforming process). In [152] and [47], block matching was performed on the RF images. Both studies showed that reasonable

cardiac deformations could be recovered. However, RF block matching can only handle small deformations [47]. The current 3D echocardiographic commercial systems have a relatively low temporal resolution, leading therefore to relatively large deformations between subsequent frames. The applications of RF tracking to clinics remains an issue. Consequently, it still remains unclear whether RF speckle tracking is preferable to the conventional B-mode tracking for this particular task [7].

Optical flow

Intensity-based optical flow : Optical flow has been applied for contour tracking in ultrasound [70]. By assuming that the intensity of a particular point in a moving pattern does not change over the image sequence, local motion vectors are solved either locally on independent image patches or globally using spatial smoothness terms as regularization. The demons algorithms represent a computationally efficient simplification of the optical flow problem [247, 257]. Somphone *et al.* developed a sparse implementation of the demons algorithms [247].

Phase-based optical flow : Optical flow was applied not only to B-mode intensities but also to phase images [3, 8]. Alessandrini *et al.* chose to perform the optical flow tracking on monogenic phases computed from the ultrasound intensity images [3]. They showed that the displacement estimated from the monogenic phase is robust to possible variations of the local image energy. Large deformations were coped with by recursively applying the estimation in a pyramidal refinement scheme.

Elastic image registration

Elastic image registration provides an alternative method to compute a smooth dense deformation field from 3D cardiac ultrasound image sequences [55]. Cardiac motion fields were represented efficiently by a set of B-spline basis functions. Diffeomorphic free-form deformation models were then used for registering echocardiographic images in time. Recently in [104], an elastic registration model was proposed to register the ultrasound images resampled in an anatomical space. The ultrasound images were first resampled to an anatomical domain of the heart following the radial, longitudinal and circumferential directions. The resampled images were then registered by a classic free-form deformation algorithm which allowed the further derivation of local motion and strain fields.

Prior model-based tracking

All of the methods mentioned above tracked the cardiac motion without using any prior information. In fact, priors can be embedded in statistical or physical forms in order to improve the motion tracking accuracy [143, 192]. For instance, Leung *et al.* used a statistical model of cardiac motion for regularizing optical flow tracking results [143]. Similarly, in [192], the authors used a transversely isotropic linear elastic model for regularizing shape-based tracking correspondences.

3 State-of-the-art in cardiac simulation

3.1 Cine MR image simulation

Tobon-Gomez *et al.* and Wissmann *et al.* simulated cardiac cine MR image sequences based on the XCAT phantom [237, 259, 285]. In [259], a tissue segmentation (or labeling) was performed where different tissue types were associated with relevant T1, T2 and proton density (PD) values

obtained from the literature. The static T1, T2 and PD maps at end-diastole (ED) were then warped by cardiac motion obtained from the XCAT phantom. Finally, at each time frame, an MR simulator was used to generate the cine MR images. Wissmann *et al.* directly simulated image intensities by signal equations instead of using an MR simulator [285]. They claim that this choice facilitates the simulation of sampling strategies, trajectory optimization and post-processing methods [285]. Haddad *et al.* and Glatard *et al.* also synthesized 3D+t cardiac cine MR images [94, 99]. They adopted a similar approach to that presented in [259] by combining a beating heart model with an MR simulator. The main difference is in the heart model, which is extracted from real acquisitions as described in [98]. Recently, Prakosa *et al.* developed a pipeline that combines an electromechanical model (E/M) of the heart with template acquisitions for generating realistic cine MR data [205]. The E/M model has the benefit of relating the contraction law with its biophysical causes which allows for a more realistic incorporation of physiological and pathological conditions. The template cine MR sequence was registered to the first frame, yielding a "frozen" sequence. This sequence was then warped with the simulation results to create synthetic images.

3.2 Tagged MR image simulation

Compared to cine MR, little work has been done in simulating cardiac tagged MR. Crum *et al.* simulated the LV in short axis slices [57]. They modeled the corresponding anatomy using a simple ring shape. Using a motion directly computed from a real cine MR sequence, the authors warped the initial simulated ED image to the rest of the sequence. The tagging pattern was simulated by applying a sinusoidal modulation function in the spatial domain. Later in [58], Crum *et al.* improved the generation of tag intensity profiles by using a frequency-domain model. Similarly, Waks *et al.* used a prolate sphere to mimic the LV geometry and a 13-parameter kinematic motion model [278]. The model parameters were determined by a least-squares fit to the displacements of the implanted markers tracked from a dog heart acquisition. Sermesant *et al.* segmented myocardium from a real tagged MR image and further added tag lines to the binary mask [238]. Finally, this image was warped by cardiac motions generated by the E/M model. Clarysse *et al.* warped a real short-axis tagged MR image by a simple kinematic mode-based heart motion model [50]. However, the motion model was too simplistic to represent the complexity of true heart motion. Similarly, in our previous work presented in [300], a normal cardiac motion pattern was first simulated by the E/M model, and a real acquisition was further deformed for assigning realistic image intensities. However, only one tagged MR sequence of a healthy heart was simulated.

3.3 Echocardiography image simulation

Existing simulation techniques mostly combine two elements : one model for anatomy and motion of the myocardium and one ultrasound simulator to mimic the image formation process. In this context considerable efforts have been spent in order to have the most realistic models for anatomy and motion and the most physically sound and computationally effective ultrasound simulations.

Preliminary work making use of synthetic cardiac US sequences include the ones in [20] and [250], where an annular shape was adopted to represent a LV short-axis cross section in combination with a radial motion model to account for contraction and expansion. Since the introduction of 3D scanners, modeling the whole muscle in 3D has received considerable attention. For example, a truncated prolate spheroid was employed in [79] to model the LV

shape in 3D with a complex kinematic model to rule the motion pattern. In [72] a truncated ellipsoid was adopted for both ventricles to which an electromechanical model of cardiac motion was applied [231]. Finally, in [54], the geometry was obtained through segmentation of cine MR images, while the Bestel-Clement-Sorine (BCS) electromechanical model [159] was used for computing deformation through the cardiac cycle. As compared to kinematic models, electromechanical ones have the benefit of directly relating the contraction law with its biophysical causes which allows for a more realistic incorporation of physiological and pathological conditions.

Regarding ultrasound simulators, FieldII [117], [115] is commonly considered as the state-of-the-art for linear acoustics : it provides a library for the calculation of pressure fields from arbitrarily shaped and apodized transducers. Nevertheless the computational burden associated to FieldII makes it unfeasible for simulating large data-sets, especially in 3D. In this context, COLE [88] was recently introduced as a fast alternative for generating 3D ultrasound sequences : it accelerates the convolution of a 3D point spread function by multiple 1D convolutions while allowing the integration of various simulated or measured beam profiles as a lookup table.

Ultrasound simulators typically model the tissue response as a collection of point scattering centers. To account for the different acoustic impedance between the myocardium and the blood pool all aforementioned frameworks proceed in the following binary fashion : high amplitude is assigned to scatterers inside the muscle and a low amplitude to the ones outside. As a consequence, the level of realism of the resulting ultrasound images remains highly unsatisfactory : surrounding structures such as papillary muscles and heart valves are neglected as well as typical artifacts such as reverberations, clutter noise, signal dropout and local intensity variations due to changing cardiac fiber orientation (see Fig. 1(a)). All these elements are critical as they have a major impact on the performance of algorithms for motion/deformation estimation.

Two solutions represented a considerable leap forward in this scenario [206], [8]. In those works, instead of a simple binary mask, a real ultrasound recording was used as a template to obtain realistic speckle textures. By doing so, surrounding structures and ultrasound artifacts were directly transferred from the template to the simulation (cf. Fig. 1(b) and (c)). However, both techniques suffer from intrinsic limitations. In [206] the synthetic sequence was obtained by warping the B-mode pixel data of the template acquisition according to the benchmark motion field obtained from a BCS model [159]. Such an approach unavoidably introduced unrealistic warping artifacts in the simulated speckle texture (cf. Fig. 1(b)). In particular, the framework enforced an unrealistically high temporal speckle correlation which could mislead the performance assessment of speckle tracking algorithms. To avoid this, in [8] an ultrasound simulator (FieldII) was used to handle the generation of the synthetic images independently. Hereto, the template sequence was used to compute the amplitude distribution of the scatter map fed to the ultrasound simulator. As compared to [206], the approach allowed generating considerably more realistic speckle images (cf. Fig. 1(c)). The synthetic motion was then obtained by applying the motion field estimated from the template sequence using the spatio-temporal elastic registration technique in [140]. This represented a main limitation since *i*) no control on the simulated motion was possible, *ii*) motion estimation errors in the template sequence were directly reflected on the simulated motion and *iii*) the benchmark displacement/deformation was naturally biased towards the registration technique employed in [140] and hence unsuited to benchmark generic motion estimation algorithms. Moreover, the approach in [8] was proposed only for 2D ultrasound.

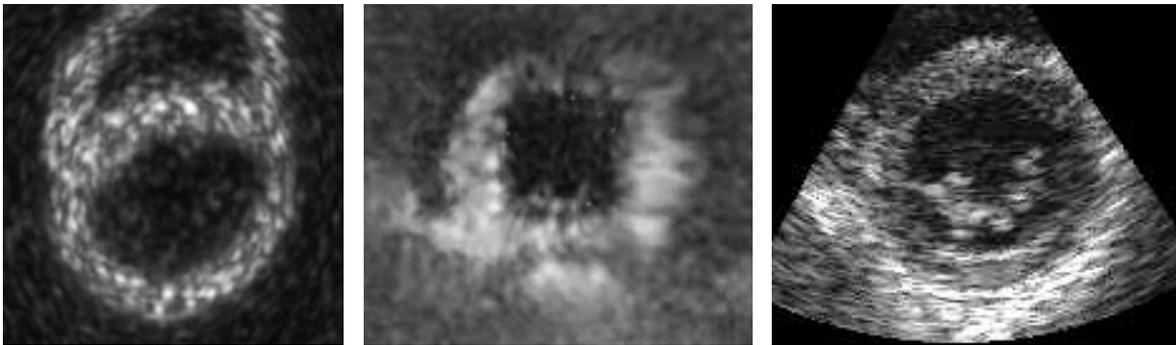


FIGURE 1 – Synthetic echocardiographic images used in [54] (a), [206] (b) and [8] (c). All techniques present intrinsic limitations which limit their usability as benchmarking tools. The solution in (a) employs a simplistic speckle model (*i.e.* high speckle intensity inside the muscle and low intensity outside) leading to unrealistic binary-like images. In (b) the simulated cine-loop is obtained by warping the B-mode pixel images of a real recording, thus producing texture warping artifacts. In (c) the presence of warping artifacts is avoided by using an ultrasound simulator. Yet, the reference motion field is computed by applying an elastic registration technique to the real recording working as template and therefore *i)* no control on the simulated motion is possible, *ii)* errors of the registration algorithm are directly reflected on the simulated motion and, moreover, *iii)* the benchmark displacement is biased towards the elastic registration technique employed. The pipeline described in this paper overcomes such limitations

4 1st contribution : realistic synthetic US and MRI sequences from the same virtual patients

In this section we present a novel pipeline specifically designed to simulate realistic synthetic US and MRI sequences from the same virtual patients. This pipeline is based on an efficient use of real clinical recordings and an electromechanical (E/M) model to generate realistic texture with controlled motion. The obtained results were analyzed to assess the degree of realism. This work has been the subject of several publications [4, 5, 7, 54, 301].

4.1 Motivations

Thanks to the litterature review proposed in Sec. 3, it appears that the different simulation pipelines suffer from a lack of realism, whatever the simulated modality. Based on this observation, we developed a pipeline to generate realistic 3D cardiac US, cine and tagged MR sequences from the same virtual patient. The proposed simulation pipeline consists of three main elements : *i*) use of real clinical recordings to extract relevant information to improve the realism of the generated images ; *ii*) involvement of an E/M model of the heart [158] to generate healthy and pathological synthetic motion field used as reference ; *iii*) use of physical simulation environments to model image formation with the possibility to introduce complex effects such as tag fading for tagged MR or speckle decorrelation for ultrasound. The proposed extended pipeline was used to create a database of 18 virtual patients including healthy and various pathological cases, *i.e.* ischemia and dyssynchrony. For each patient, benchmark sequences of 3D US, cine and tagged MR were generated. We generated a total of 90 synthetic sequences representing 2700 image volumes. The synthetic sequences along with the corresponding reference motion fields are made publicly available via an open-access database (<http://humanheart-project.creatis.insa-lyon.fr/multimodalityStraus.html>).

4.2 Overview

A schematic view of the developed pipeline is shown in Fig. 2. Each block specifies the number of the subsection where it is described in detail. Briefly, real acquisitions (cine/tagged MR) are used as template sequences (Sec. 4.3) for deriving realistic textures. The first step is to segment the right ventricle and the left ventricle on the first frame of cine MR template sequence (Sec. 4.4.1). The E/M model is then applied to the segmented 3D geometry to simulate one cardiac cycle (Sec. 4.5.1). For each frame of the simulated sequence, US, cine and/or tagged MR images are generated from a physical simulator which takes relaxation times T_1 and T_2 and proton density maps as inputs for MR simulations (Sec. 4.4) and a scattering map as input for US simulations (Sec. 4.7.1). To achieve realistic texture, proton density maps (respectively scattering maps) are derived from the MR (respectively US) template recordings. This requires a dedicated spatio-temporal alignment between the simulations and the template recordings (Sec. 4.6). As such, the pipeline will generate synthetic US, cine and tagged MR sequences with a fully controlled deformation field and with a texture visually similar to the one of the corresponding template recording.

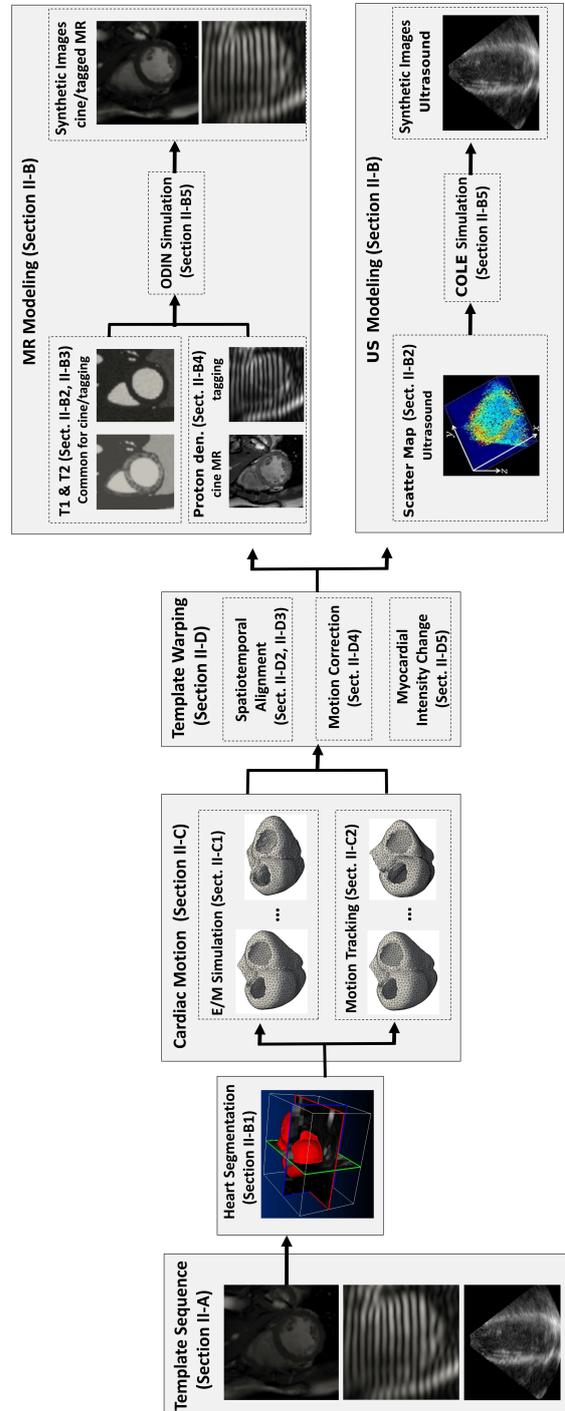


FIGURE 2 – Proposed pipeline for the simulation of realistic cardiac US, cine and tagged MR sequences.

4.3 Template image sequences

3D template recordings used in this study come from the open access database provided in [258]. For each volunteer, one US, one cine MR and three tagged MR sequences were acquired. The MR datasets were acquired using a 3T Philips Achieva System (Philips Healthcare, Best, The Netherlands). The MR sequences processed by the participants were cine steady state free precession (SSFP) [236] and 3D tagged MR (3DTAG) [230]. Cine MR images were acquired during breath-holds of approximately 15 seconds and were gated to the vector ECG. 3DTAG datasets were obtained with three sequential breath-hold acquisitions in each orthogonal directions. Since MR images were acquired in the same patient coordinate system, cine and tagged MR sequences are naturally aligned in space. A set of 24 manual landmarks identifying the same keypoints in all the modalities allows the derivation of a simple rigid transformation to pass from MR to US coordinates. These landmarks are evenly distributed over the LV, with 8 landmarks at each of the three ventricular levels (basal, midventricular, apical) [258]. All the acquisitions represent one cardiac cycle. The US template has a spatial resolution of $0.82 \times 0.83 \times 0.73 \text{ mm}^3$ while the cine and tagged MR templates have spatial resolutions of $1.25 \times 1.25 \times 8.08 \text{ mm}^3$ and $0.96 \times 0.96 \times 7.71 \text{ mm}^3$.

To make the explanation of the full pipeline easier, we first describe the MR simulation from Sec. 4.4 to Sec. 4.6. Based on the derived formulation, we then give details on the ultrasound simulation (Sec. 4.7).

4.4 MR modeling

We used the ODIN simulator [121] to generate both cine and tagged MR images. ODIN is a physical simulator which allows either the selection or the design of specific MR sequences. To simulate one 3D image, ODIN requires T_1 , T_2 and proton density maps as inputs. In this study, we used the same T_1 and T_2 maps to generate both cine and tagged MR sequences for each virtual patient, only the proton density maps were computed separately for each modality and for each time instant of the simulated sequence. The first step of the proposed MR pipeline is the computation of the initial T_1 and T_2 maps for the first frame. We first describe how these initial maps were generated. We then give details on the way they were extended to all time frames.

4.4.1 Heart segmentation

Because we assign different T_1/T_2 values per structure, we first need to segment the initial template image into different classes, *i.e.* myocardium, blood, lung and air. Since cine MR provides the best image contrast and since all the template modalities can easily be mapped, it represents a natural choice to extract the heart geometry. We first used the VP2HF segmentation pipeline proposed in [96] to segment the myocardium and the underlying blood pools. We then segmented the lung and the air outside the body thanks to simple mathematical morphology operations. Finally, the remaining unlabeled pixels were classified by thresholding : high intensity pixels belonging to blood and the rest to muscles. Fig. 3(b) shows a typical segmentation result obtained by using such procedure.

4.4.2 Initial T_1 and T_2 maps

We assigned T_1 and T_2 maps using published reference mean and standard deviation values (denoted as μ and σ) available from [259]. At each pixel location, T_1 and T_2 values were randomly

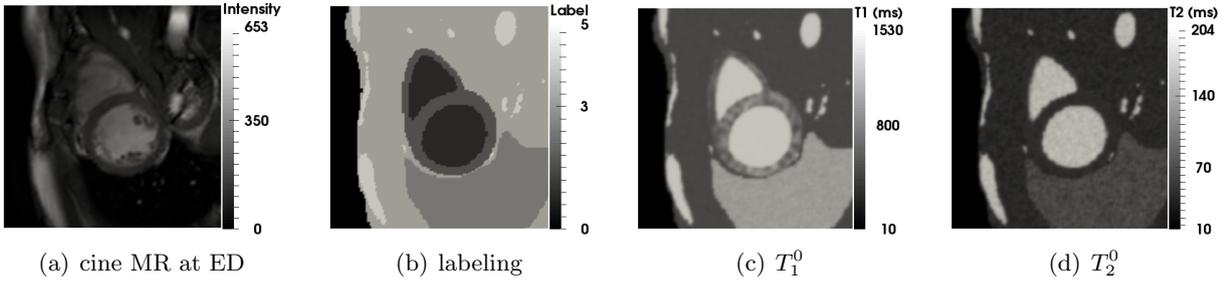


FIGURE 3 – Illustration of the different steps involved to simulate the initial T_1 and T_2 maps relative to the first frame (ED).

sampled from Gaussian distributions, yielding initial static maps denoted as T_1^0 and T_2^0 (see Fig. 3(c) and 3(d)). The relevant u and σ values used in this study are listed in Table 1. Moreover, since the myocardium is our region of interest, a dedicated strategy was applied to better control its underlying texture. Indeed, since tagged MR intensity changes contain strong T_1 information, it is possible to deduce relevant local and structural T_1 values from image intensities. In particular, temporal intensity changes of the CSPAMM tagged MR sequence can be described as follows [85] [280] :

$$I_{tag}^n = I_{tag}^0 \exp \left\{ -\frac{n\Delta t}{T_1} \right\} \frac{\sin(\alpha_n)}{\sin(\alpha_0)} \prod_{j=0}^{n-1} \cos(\alpha_j), \quad \forall n \geq 1 \quad (1)$$

where Δt is the temporal resolution, n is the frame index and I_{tag}^n is the tagging intensities of the material point at frame n . α_n corresponds to the RF flip angle at frame n and can be easily computed from the method described in the appendix of [301]. Eq. (1) thus gives access to T_1 information from the intensity changes of the tagged MR template sequence. In order to accurately estimate the T_1 map of the myocardium, the following procedure was repeated for all pixels belonging to this region. For a given point at ED, we tracked it through the entire tagged MR template sequence (this step is later described in Section 4.5.2). This gives a sequence of intensities I_{tag}^n for all time instants. The corresponding T_1 value was then estimated from linear least square minimization problem :

$$\arg \min_{T_1} \sum_n \left(\log \left(\frac{I_{tag}^n \sin(\alpha_0)}{I_{tag}^0 \sin(\alpha_n) \prod_{j=0}^{n-1} \cos(\alpha_j)} \right) T_1 + n\Delta t \right)^2 \quad (2)$$

Since there are three template sequences of tagged MR images per subject (Section 4.3), we derived three different T_1 values for the same myocardium point. A median filtering was then applied for removing possible outliers, yielding the final T_1^0 myocardial map as shown in Fig. 3(c).

4.4.3 Dynamic T_1 and T_2 maps

Since each frame of the simulated sequences was generated independently, the initial T_1^0 and T_2^0 maps previously described had to be extended over time. This was achieved by a dedicated warping strategy described in Section 4.6. The corresponding dynamic maps are denoted as T_1^k and T_2^k , k indexing the simulation frame.

TABLE 1 – T_1 and T_2 reference values attached to each label of the segmented cine MR image.

Label	Class	T1 (ms)		T2 (ms)	
		u	σ	u	σ
0	Air	0	0	0	0
1	Blood	1516	21	189	26
2	Myocardium	982	46	54	12
3	Lung	1199	117	79	29
4	Other (low)	549	52	49	20
5	Other (high)	1516	21	189	26

4.4.4 Proton density maps

For each simulation frame, we further computed the proton density map from the corresponding template recordings. This ensures the generation of a texture visually similar to the one of the underlying template. To this aim, one needs to spatially and temporally align each frame of the simulated sequence with its counterpart in the template sequence. This was achieved by a dedicated warping strategy detailed in Section 4.6. The outputs of this step correspond to a set of warped templates denoted as \tilde{I}_{cine}^k and \tilde{I}_{tag}^k and aligned with each simulation frame k .

Cardiac cine MR images were acquired using the balanced steady state free precession (bSSFP) MR sequence [236]. The proton density map ρ_{cine}^k of frame k can then be computed from the corresponding warped template [236] :

$$\rho_{cine}^k = \tilde{I}_{cine}^k \sqrt{T_1^k / T_2^k} \quad (3)$$

For tagged MR, the characteristics of the CSPAMM sequence provide the following ρ_{tag}^k map [280] :

$$\rho_{tag}^k = \tilde{I}_{tag}^k \exp^{\frac{TE}{T_2^*}} / \sin(\alpha_k^s) \quad (4)$$

where α_k^s corresponds to the simulated RF flip angle, where TE is the echo time here set to 21 *ms*. T_2^* corresponds to the effective spin-spin relaxation times and are here approximated by the relaxation time T_2 [97]. We thus assume $T_2^* = T_2^k$ in the sequel.

4.4.5 ODIN simulation

Each frame k of the simulated sequences was independently generated from ODIN by using the T_1^k , T_2^k and ρ_{mod}^k maps derived above (with $mod = cine$ or tag) as inputs. For the sake of simplicity, we decided to use the same resolution for those maps and the generated output MR images. Moreover, since we wanted to reproduce as close as possible the template recordings, we used the image properties (origin, spacing, size and axis orientations) of the template cine and tagged MR as reference. This defines the pixel positions of the simulated images for each modality. The T_1^k , T_2^k and ρ_{mod}^k maps were thus interpolated to those pixel positions before being passed on to ODIN. We used respectively the *odinisp* and *odinepi* MR sequences available under ODIN for simulating cine and tagged MR images. Indeed these sequences are consistent with the corresponding real acquisition protocol, *i.e.* the bSSFP for cine sequence [258] and the EPI for tagged sequence [230].

4.5 Cardiac motion

As commented in Section 4.4.1, the first frame of the template cine MR sequence was segmented into different classes including myocardium region (LV and RV). From this mask, a volumetric mesh (both LV and RV) was generated using [5] and propagated to all further simulation and template image time frames as described hereinafter.

4.5.1 E/M simulation

From the myocardial mask of the first frame, a tetrahedral mesh was generated using the CGAL¹² meshing software [2] and passed to the E/M simulator which generated as output a sequence of volumetric meshes \mathcal{S}_k . The E/M simulator applies the Bestel-Clement-Sorine model which estimates the electrical activation and the resulting mechanical contraction [158] and is implemented in the SOFA framework¹³. This model was chosen for its realistic properties and it complies with the four cardiac phases (isovolumic contraction, ejection, passive and active filling). Cardiac fiber orientations were estimated with a rule-based method (elevation angle between -70° to $+70^\circ$) in order to model the E/M anisotropy. It has already shown its efficiency in the generation of realistic US sequences [5]. A number of biophysical parameters such as the local myocardial contractility, stiffness and conduction governs the E/M equations. By varying those parameters, both normal and pathological cardiac motion fields can be simulated. For each patient, we generated in this study one healthy, one dyssynchrony (LBBB) and four realistic ischemic hearts. For the dyssynchrony, only the right endocardium was initially activated. For the ischemia, varying scar locations with random and realistic shapes were simulated [75] in parts of the LV where arterial occlusion can occur : one for the Left Anterior Descending artery (LAD), one for the Left Circumflex (LCX), and two for Right Coronary Artery (RCA). Corresponding figures could be found in the supplementary materials of our paper [301].

4.5.2 Motion Tracking

Since the cine and tagged MR were acquired under the same coordinate system, we used the tagged MR sequence to perform the tracking. The LV and RV of the first E/M model mesh \mathcal{S}_0 (the E/M model and the template are naturally aligned at the first frame) were extracted and processed separately by two state-of-the-art algorithms in order to maximize the tracking accuracy. The LV border was tracked using the HarpAR algorithm given its good ability in estimating myocardial deformations, especially the radial strain [299]. The sparse Demons algorithm was then selected to perform RV tracking given its good tracking accuracy and its low computational cost [247]. The output of this procedure corresponds to a unique sequence of segmented meshes \mathcal{R}_t attached to both cine and tagged MR.

4.6 Template Warping

As described in Section 4.4.4, a spatio-temporal alignment between the template sequences and each simulation frame is needed to derivate the proton density maps. We developed a novel warping framework to tackle this problem. Similar to [5], we used two different strategies to deal with the myocardium and the surrounding structures (named as background in the sequel), however we introduce the following innovative aspects : *i*) we developed a dedicated strategy to ensure smooth transition at the interface between the myocardium and the background ; *ii*)

12. the open source Computational Geometry Algorithms Library (CGAL) is available at www.cgal.org.

13. <https://www.sofa-framework.org/>

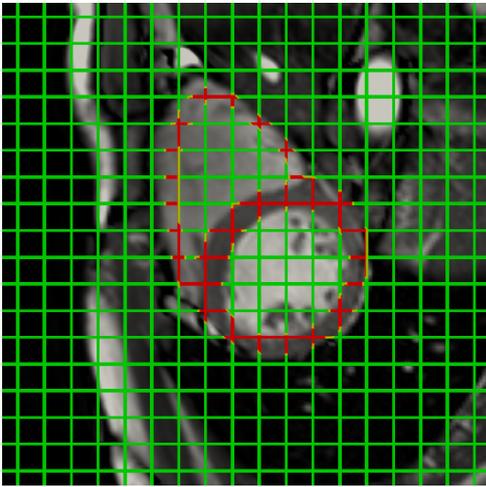
we introduced a model to handle temporal myocardial intensity changes over the cardiac cycle. Each step of the proposed procedure is detailed hereunder.

4.6.1 Generic transformation model

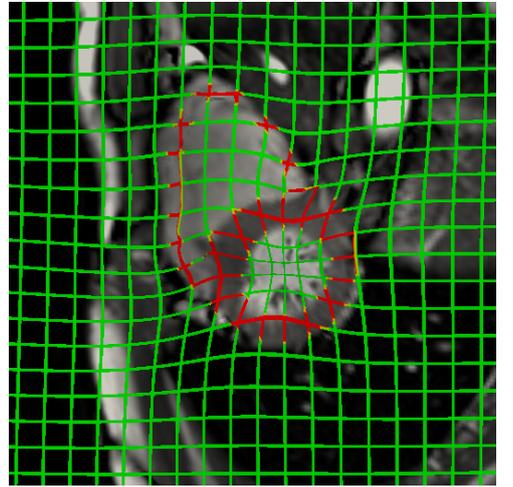
Similar to [5], we used myocardial landmark points extracted from the two sequences of meshes \mathcal{S}_k and \mathcal{R}_t to compute displacement fields. In this study, we added static background landmarks (denoted as \mathcal{B}) to avoid non-physiological movement of the structures surrounding the heart muscle. These landmarks were detected from the cine MR template sequence. In particular, for each pixel position, we computed the variation of cine MR intensity over the whole sequence and we kept the points with the smallest variation and with a minimal spacing of 5 mm. For each simulation, the number of static background landmarks was fixed to 200.

The goal of the introduced generic model was to compute a global transformation to efficiently match a source space (represented by the heart mesh \mathcal{M}_0) with a target space (represented by the heart mesh \mathcal{M}_1) where \mathcal{M}_0 and \mathcal{M}_1 share the same number of points. In both spaces, we assumed the presence of the same static background landmarks \mathcal{B} . The transformation maps one pixel \mathbf{x} of the source space to a corresponding position \mathbf{y} in the target space. As stated above, different transformation strategies were applied to the myocardium and the background.

Regarding the myocardial points, since they all lie inside the volumetric mesh \mathcal{M}_0 , the corresponding transformation was directly computed from the two volumetric meshes $\mathcal{M}_0 \rightarrow \mathcal{M}_1$. The barycentric coordinates of \mathbf{x} in mesh \mathcal{M}_0 were evaluated. From these barycentric coordinates and the node positions of the target mesh \mathcal{M}_1 , we computed the absolute coordinates. This gave the matched position \mathbf{y} . This procedure is denoted as $\mathbf{y} = MESH_{\{\mathcal{M}_0 \rightarrow \mathcal{M}_1\}}(\mathbf{x})$.



(a) initial cine MR image at ED



(b) image deformed to ES

FIGURE 4 – Illustration of the use of the generic transformation model. The ED cine MR image was warped to ES by the generic transformation. The grid superimposed on the image shows the corresponding deformation. The myocardium and the background are shown in red and green respectively. The generic transformation was computed from the ED/ES simulation meshes and the static background points.

Concerning the background points, we used the Thin Plate Spline (TPS) algorithm to model

the underlying transformation. We randomly selected a fraction of mesh points (800 points in all the simulations) which, together with the static landmarks, were used to parameterize the TPS algorithm. This procedure is denoted as $\mathbf{y} = TPS_{\{\mathcal{M}_0, \mathcal{B} \rightarrow \mathcal{M}_1, \mathcal{B}\}}(\mathbf{x})$. The overall transformation can thus be written as :

$$\begin{aligned} & \mathcal{T}_{\{\mathcal{M}_0, \mathcal{B} \rightarrow \mathcal{M}_1, \mathcal{B}\}}(\mathbf{x}) \\ &= \begin{cases} MESH_{\{\mathcal{M}_0 \rightarrow \mathcal{M}_1\}}(\mathbf{x}), & \text{if } \mathbf{x} \in \text{Myocardium} \\ TPS_{\{\mathcal{M}_0, \mathcal{B} \rightarrow \mathcal{M}_1, \mathcal{B}\}}(\mathbf{x}), & \text{otherwise} \end{cases} \end{aligned} \quad (5)$$

Fig. 4 shows an example of the use of the generic transformation defined by Eq. (5) to derive a deformation map between two time instances.

4.6.2 Temporal alignment

Temporal alignment was performed by linearly stretching/shrinking the time axis of the template recording in order to match relevant cardiac events used as temporal landmarks in the simulation sequence. In this paper, ED and ES time events were chosen as landmarks. Each frame k of the simulation was matched to a cardiac time (ms) in the template sequence by the ϕ operator defined as :

$$\phi(k) = \begin{cases} \frac{t^{es}}{k^{es}}k, & \text{if } k \leq k^{es} \\ t^{es} + \frac{k-k^{es}}{K-k^{es}}(t^{max} - t^{es}), & \text{otherwise.} \end{cases} \quad (6)$$

K is the number of simulated frames and k^{es} corresponds to the ES frame index. Similarly, t^{max} is the template cardiac cycle while t^{es} refers to the template ES timing. Both the template cine and tagged MR sequences were resampled to cardiac timings $\phi(k)$, with $k \in [0, K - 1]$. In each template sequence, the two frames with the cardiac times closest to $\phi(k)$ were found. These two images were then linearly interpolated to $\phi(k)$, producing template sequences temporally aligned with the simulation frames and denoted as \hat{I}_{mod}^k .

4.6.3 Spatial alignment

After temporal synchronization, spatial alignment was needed to align the resampled images \hat{I}_{mod}^k with the corresponding E/M geometry \mathcal{S}_k . To this aim, the generic transformation introduced in Sec. 4.6.1 was used, yielding to the following expression :

$$\hat{I}_{mod}^k = \hat{I}_{mod}^k \circ \mathcal{T}_{\{\mathcal{R}_{\phi(k)}, \mathcal{B} \rightarrow \mathcal{S}_k, \mathcal{B}\}} \quad (7)$$

where \hat{I}_{mod}^k is the warped template image aligned both in time and space with the corresponding simulation mesh \mathcal{S}_k . The full spatio-temporal alignment strategy is given in Fig. 5.

4.6.4 Motion Correction

The warped template sequences \hat{I}_{mod}^k rely on motion tracking which may introduce artifacts in the synthetic images by adding an apparent residual motion to the true motion given by the E/M model. Contrary to [205], we introduced in this study a dedicated strategy that corrects the myocardial motion in the warped sequences. To ensure a coherence with the E/M model, the myocardial intensities of each frame of the simulated sequences were sampled from the

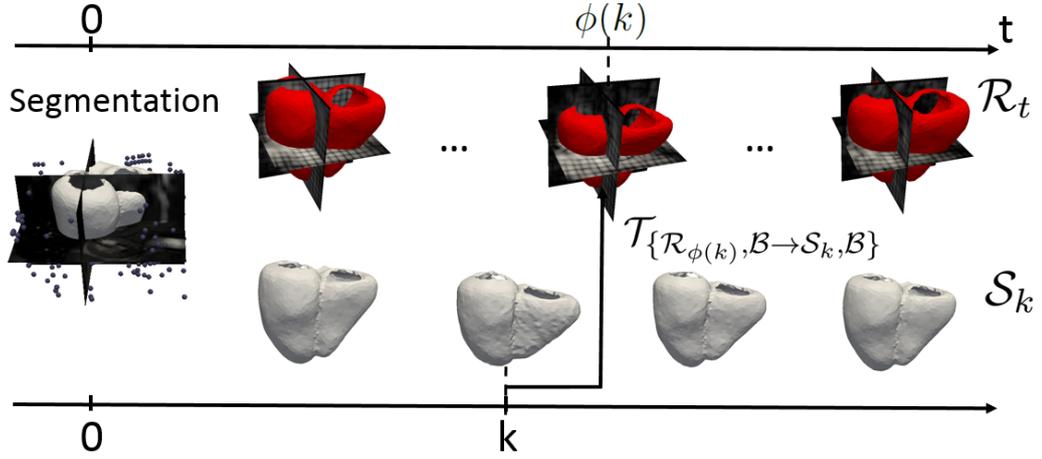


FIGURE 5 – Illustration of the temporal and spatial alignments described in Section 4.6.2 and 4.6.3 respectively. The myocardium is segmented at ED. The static background points are shown as black dots surrounding the mesh. The segmentation mesh is first tracked in the template sequence, yielding the tracking meshes \mathcal{R}_t shown in red (the upper row). Then, the E/M model is exploited for simulating myocardial deformations starting from the same segmentation mesh, leading to another sequence of meshes \mathcal{S}_k shown in the lower row. Each simulation frame k is mapped to a cardiac time $\phi(k)$ of the template through the temporal alignment. The spatial alignment $\mathcal{T}_{\{\mathcal{R}_{\phi(k)}, \mathcal{B} \rightarrow \mathcal{S}_k, \mathcal{B}\}}$ then register the template image to the corresponding simulation mesh \mathcal{S}_k .

first template images. This was achieved by deforming I_{mod}^0 to all simulation frames, yielding a complementary warped template sequence \bar{I}_{mod}^k expressed as :

$$\bar{I}_{mod}^k = I_{mod}^0 \circ \mathcal{T}_{\{\mathcal{S}_0, \mathcal{B} \rightarrow \mathcal{S}_k, \mathcal{B}\}} \quad (8)$$

This warped sequence was used to provide myocardial intensities while the warped sequence \hat{I}_{mod}^k was used to get background intensities. In order to ensure a smooth interface between these two regions, we computed a Gaussian shaped weighting function centered at the myocardium for each simulation frame. Since most tracking algorithms aim to quantify the LV myocardium deformations, we only correct in this work the motion attached to this structure. For each simulation mesh \mathcal{S}_k we extracted the LV surface and further converted it to an implicit function denoted as Γ_k . This function equals to 0 at the endo- and epicardium borders, is negative inside the myocardium and positive outside as is shown Fig. 6. Its minimal value is denoted as τ_k^{min} . The weighting function was then defined as follows :

$$w_k(\mathbf{x}) = \exp \left\{ -\frac{(\Gamma_k(\mathbf{x}) - \tau_k^{min})^2}{2\sigma^2} \right\} \quad (9)$$

where σ controls the transmural profile of the weighting function and was set to 5 in all the simulations. The warped template sequences with myocardial motion correction were then computed as follows :

$$\mathcal{IM}_{mod}^k = w_k \bar{I}_{mod}^k + (1 - w_k) \hat{I}_{mod}^k \quad (10)$$

Eq. (10) reveals that most of the myocardial intensities are assigned from the first template frame, ensuring coherent myocardial motion with the E/M model. The background intensities

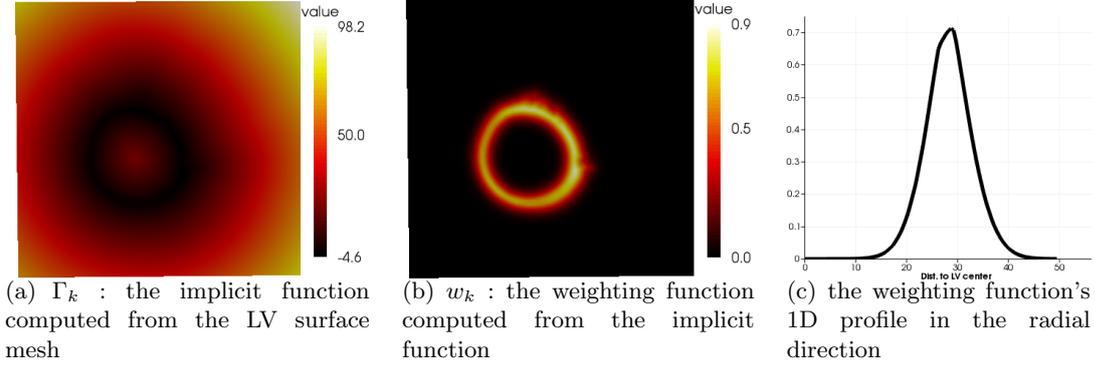


FIGURE 6 – The computation of the Gaussian shaped weighting function.

are principally retrieved from the spatiotemporally aligned template sequence, preserving the realistic nature of the surrounding structures.

4.6.5 Myocardial intensity changes

The warped template sequences \mathcal{IM}_{mod}^k involved constant myocardium intensities over the cardiac cycle, which is too simplistic. We thus developed dedicated strategies for each modality to integrate myocardial temporal intensity changes in the simulated sequences.

Regarding tagged MR, Eq. (1) reveals that tagging intensity changes depend on T_1 relaxation and the RF flip angle. We thus used the static T_1^0 map together with the simulated RF flip angles α_k^s (details about its computation can be found in the appendix of our paper [301]) to reproduce myocardial intensity variations for each simulated frame k as follows :

$$\Delta \mathcal{IM}_{tag}^k = I_{tag}^0 \left(\exp \left\{ -\frac{k \Delta t^s}{T_1^0} \right\} \frac{\sin(\alpha_k^s)}{\sin(\alpha_0^s)} \prod_{j=0}^{k-1} \cos(\alpha_j^s) - 1 \right) \quad (11)$$

where Δt^s corresponds to the simulated temporal resolution, while $\Delta \mathcal{IM}_{tag}^k$ corresponds to a non-warped sequence characterizing the temporal intensity changes at each pixel location.

Concerning cine MR, it can be observed that myocardial intensities increase from ED to ES and gradually decrease to their initial values at the end of the cardiac cycle. We assume that the intensity varies linearly during these two phases. The intensity change was defined as a percentage of the initial intensity. This percentage equals to 0 at ED, ξ at ES and changes linearly in between as follows :

$$\lambda_\xi(t) = \begin{cases} \frac{t}{t^{es}} \xi, & \text{if } t \leq t^{es} \\ \frac{t^{max} - t}{t^{max} - t^{es}} \xi, & \text{otherwise} \end{cases} \quad (12)$$

where t^{es} and t^{max} were introduced in Section 4.6.2. The ξ parameter was automatically computed from the template cine MR sequence for all myocardial points. In particular, for a given myocardial point \mathbf{x}_i , we first retrieved the corresponding intensity values for all time instants thanks to the protocol described in Section 4.5.2. These intensities are denoted as $I_{cine}^n(\mathbf{x}_i)$ in the sequel,

where n corresponds to the cardiac cine MR frame instant. Each ξ value related to a myocardial point \mathbf{x}_i was computed by fitting a linear model :

$$\arg \min_{\xi} \sum_n (I_{cine}^n(\mathbf{x}_i) - (1 + \lambda_{\xi}(t_n))I_{cine}^0(\mathbf{x}_i))^2 \quad (13)$$

Since Eq. (13) has a quadratic form with respect to ξ , its resolution is straightforward. For the points outside the myocardium, ξ was set to 0, meaning that their intensity does not vary over time. Finally, the intensity variations for each simulated frame k was computed as :

$$\Delta \mathcal{I}M_{cine}^k = \lambda_{\xi}(\phi(k)) \cdot I_{cine}^0 \quad (14)$$

where $\phi(k)$ is the temporal alignment introduced in Section 4.6.2 and I_{cine}^0 is the first template cine MR image.

The myocardial intensity variations $\Delta \mathcal{I}M_{mod}^k$ were finally added to the previously generated warping images $\mathcal{I}M_{mod}^k$. Since the intensity variation maps were computed in static (*i.e.* without the integration of any motion), they had to be deformed before summation. The final warped template sequences with motion correction and integrating myocardial intensity changes were computed as follows :

$$\tilde{I}_{mod}^k = \mathcal{I}M_{mod}^k + w_k (\Delta \mathcal{I}M_{mod}^k \circ \mathcal{T}_{\{S_0, \mathcal{B} \rightarrow S_k, \mathcal{B}\}}) \quad (15)$$

where w_k is the weighting function given by Eq. (9). \tilde{I}_{cine}^k and \tilde{I}_{tag}^k are therefore the final warped template cine/tagging sequences used to compute the proton density maps, as described in Section 4.4.4.

4.7 US simulation

We used COLE simulator [88] to generate US images. COLE is a fast alternative to the physical simulator such as FieldII [117]. To simulate one 3D US image, COLE requires 3D cloud of point scatterers mimicking the local tissue echogenicity for each time instant of the simulated sequence and the characteristics of the spatially variant point spread function of the simulated imaging system as inputs. The first step in the ultrasound simulation is thus the definition, at each simulation time k , of a 3D cloud of point scatterers. Each point scatterer is assigned a scattering amplitude related to the amount of energy reflected back to the probe (*i.e.* the intensity of the echo signal). Our approach consists in moving the scatterers according to the E/M simulation while sampling their amplitude from the real template recording. This will generate synthetic sequences with fully controlled cardiac motion and realistic ultrasound speckle texture.

4.7.1 Scatter Map

We call $\mathbf{x}_i^k = [x_i^k, y_i^k, z_i^k]$ and a_i^k the position and the amplitude of the i -th scatterer at simulation time k . We call $\mathcal{I}M_{us}^k(\mathbf{x}) \in [0, 255]$ the intensity at position \mathbf{x} of the warped template sequences defined in Eq. (10). Note that when non-integer spatial positions need to be accessed, the intensity value is obtained by using bi-cubic interpolation.

The first step is to generate a scatter map for the first frame $k = 0$. The coordinates of the scatterers are obtained by uniformly sampling N_{scatt} points \mathbf{x}_i^0 through the image domain. Their echogenicity is obtained by sampling the intensity of the template frame, namely $a_i^0 = F(I_{us}^0(\mathbf{x}_i^0))$,

where F is the non linear transformation applied to the intensity value in order to compensate for the log-compression traditionally performed in the ultrasound device prior to display. In particular, F is defined in such a way that :

$$20 \log_{10} [F(I(\mathbf{x})) + dB] = dB \times \frac{I(\mathbf{x})}{255}, \quad (16)$$

where dB is the desired contrast in decibel in the simulated image.

Given \mathbf{x}_i^0 , the first simulation mesh is used to distinguish myocardial scatterers (*i.e.* contained inside the mesh) from background ones (*i.e.* outside the mesh). We denote the number of myocardial and background scatterers as N_{myo} and N_{back} , respectively. The barycentric coordinates of each myocardial scatterer relative to the associated mesh tetrahedron are computed and stored.

The scatter map generation at times $k > 0$ proceeds differently for myocardial scatterers and background ones. Myocardial scatterers move anchored to the simulation meshes to mimic the motion of cardiac tissue. This is implemented by keeping their barycentric coordinates constant over time. From the barycentric coordinates and the nodes position of the simulation mesh at the considered time $k > 0$, the absolute coordinates of myocardial scatterers \mathbf{x}_i^k are then computed. As *e.g.* in [72], [54], [8], the amplitude of myocardial scatterers is kept constant over time, *i.e.* $a_i^k = a_i^0$. Such a choice of assigning scattering amplitude ensures having temporal coherency of the speckle texture patterns inside the myocardium. However, the underlying intensity can still vary with relative position changes over the cardiac cycle between the probe and the scatterers. As common in clinical scans, the apical portion of the LV epicardium was not fully visualized in the template recording (cf. Fig. 10). As such, no amplitude information was available for myocardial scatterers in that region. To avoid the appearance of holes in the simulated sequences during cardiac contraction, such scatterers were assigned an artificial amplitude equal to the average of the 50 closest scatterers inside the field of view (*i.e.* for which amplitude information was available from the image content). This solution was found sufficient to avoid artifacts.

At each simulation time $N_{back} = N_{scatt} - N_{myo}$ background scatterers are re-drawn outside the myocardium following a uniform random distribution. Their random motion accounts for the lack of speckle coherency which is normally observed in the blood pool. Background scatterers also mimic the appearance and motion of surrounding structures as valves and papillary muscles which are not described by the electromechanical model. Background scatterers are visualized by updating their amplitude at each simulation time by using the intensity from the matched position in the template recording, as defined by the spatiotemporal mapping in Sec. 4.6.4. Namely, given a background scatterer \mathbf{x}_i^k , the intensity used to compute the corresponding backscattered amplitude a_i^k is directly obtained from $\mathcal{I}\mathcal{M}_{us}^k(\mathbf{x}_i^k)$. Note that the perceived motion of background structures comes from the intensity variations and does not correspond to the TPS displacement field \mathcal{T} , which only serves to relate the two geometries. As such, while the rendered motion of surrounding structures is visually realistic, their reference displacement is not available.

4.7.2 Point Spread Function

From the 3D scatter map an ultrasound volume was simulated by convolution with the spatially variant point spread function (PSF) of the simulated imaging system. In particular COLE [88] was adopted as ultrasound simulation environment due to its good compromise

between computational efficiency and accuracy [89]. The properties of the synthetic probe were set so to match as close as possible the ones of the transducer used in the acquisition of the template. Namely, the synthetic US system was sampling at 50 MHz and equipped with a phased array transducer, which was centered at 3.3 MHz and transmitting a Gaussian pulse with a -6 dB relative bandwidth of 65%. A symmetric transverse two-way beam profile was assumed, focusing at 80 mm when transmitting and dynamically focusing on receive. The simulated images consisted of 107×80 lines in azimuth and elevation direction over an angle of 76×76 degrees, resulting in a frame rate of 30 Hz assuming parallel beam forming and ECG gating. After envelope extraction, log compression and scan conversion B-mode voxel data were available.

4.8 Result

We used in this study three volunteers from [258]. For each volunteer, we simulated one healthy and five pathological cases (Section 4.5.1), yielding an open access database composed of 18 virtual patients. For each patient, five synthetic sequences were generated (one cine MR, three tagged MR and one US). All the simulations were launched from the virtual imaging platform VIP¹⁴[94] which allows the execution of applications such as ODIN as a web service and benefits from the EGI¹⁵ computing power. The generation of one full 3D sequence took around 6 hours on VIP, knowing that it would have taken over 280 hours¹⁶

4.8.1 Image properties : qualitative assessment

The proposed generic pipeline allows the simulation of multimodality volumes visually similar to real template recordings, as illustrated by Fig. 7 to 9. In particular, Fig. 7 display the three dynamic 2D MR slices located at the basal, mid and apical regions of the LV. Fig. 8 illustrates the three channels of tagged MR simulated in this study. As for the US images, the conventional short-axis, 2-chamber and 4-chamber views are displayed in Fig. 9. Global deformation values were in clinical ranges (as detailed in Sec. 4.8.4). To give the readers a straightforward visual comparison between the template and the simulation, we show in Fig. 10 and Fig. 11 the ultrasound and MR simulations corresponding to the LCX pathology. The simulated images were shown alongside the templates. We observe that the simulations have image textures visually similar to the templates. Fig. 12 also reveals the efficiency of the Gaussian weighting function introduced in Sec. 4.6.4 to deal with the interface between the myocardium and the surrounding structures. Indeed, the proposed scheme allows a smooth transition between the different structures in this region while preserving a consistent motion inside the myocardium with the E/M model.

4.8.2 MR image properties : quantitative assessment

The level of realism of the simulated MR images was assessed quantitatively by measuring the intensity distribution inside the myocardium since it is the feature commonly exploited for the analysis of MR cardiac datasets. The myocardial histograms at ED and ES for both a simulated cine MR sequence and its corresponding template are reported in Fig. 13. The analysis first demonstrated that the computed histograms showed very similar shapes at ED. This is consistent with the fact that ED corresponds to the reference time in the proposed pipeline from which T1,

14. <https://vip.creatis.insa-lyon.fr>

15. <https://www.egi.eu>

16. estimation based on the time needed for computing one 2D slice on a personal computer with 16G of RAM and a 2.70Ghz CPU on a personal laptop.

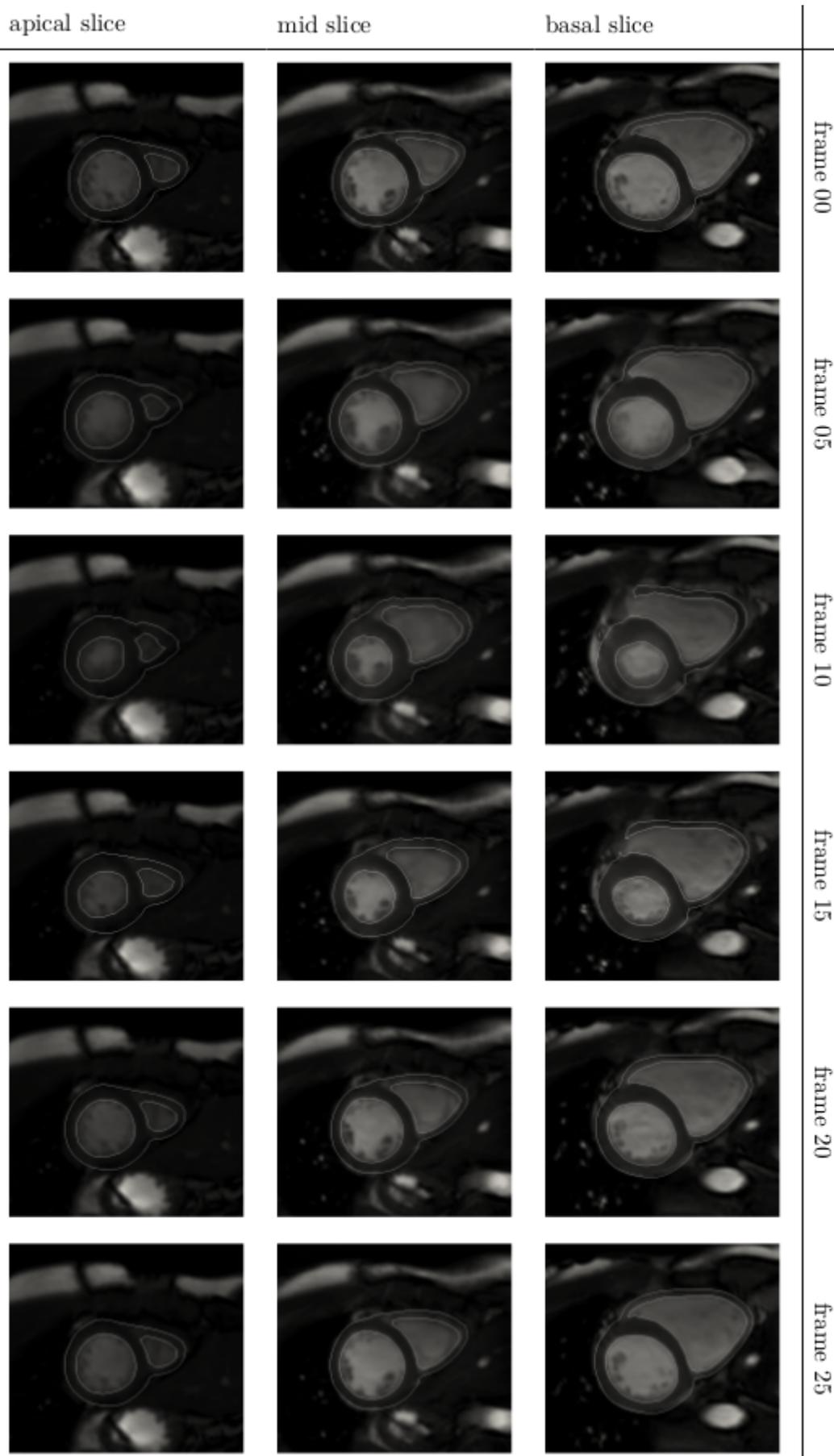


FIGURE 7 – Simulated cine MR images. Three slices located at basal, mid and apical regions of the left ventricle are shown. The mesh contour is superposed on the images.

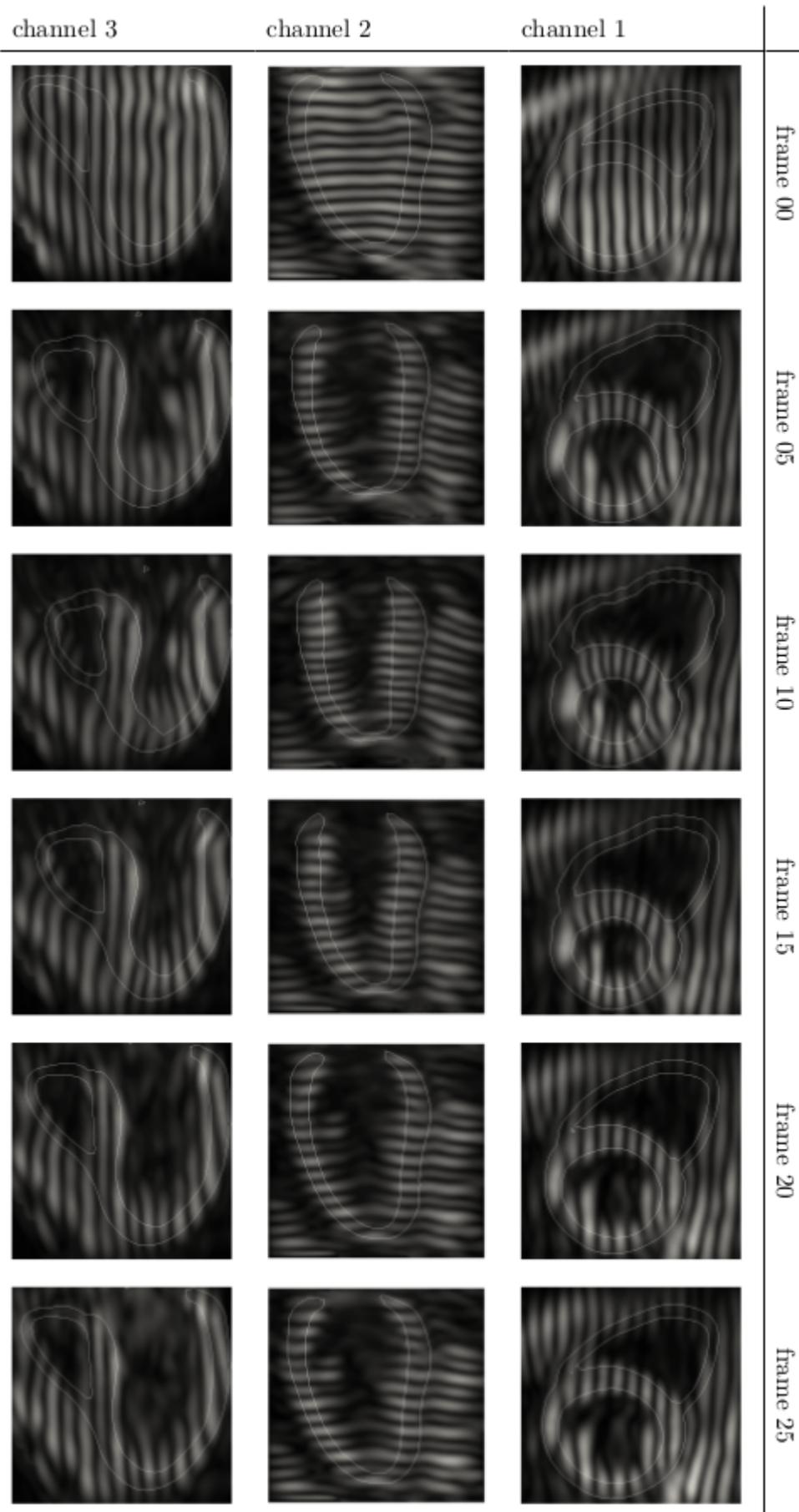


FIGURE 8 – Simulated tagged MR images. Three channels with orthogonal tag orientations are simulated. The mesh contour is superposed on the images.

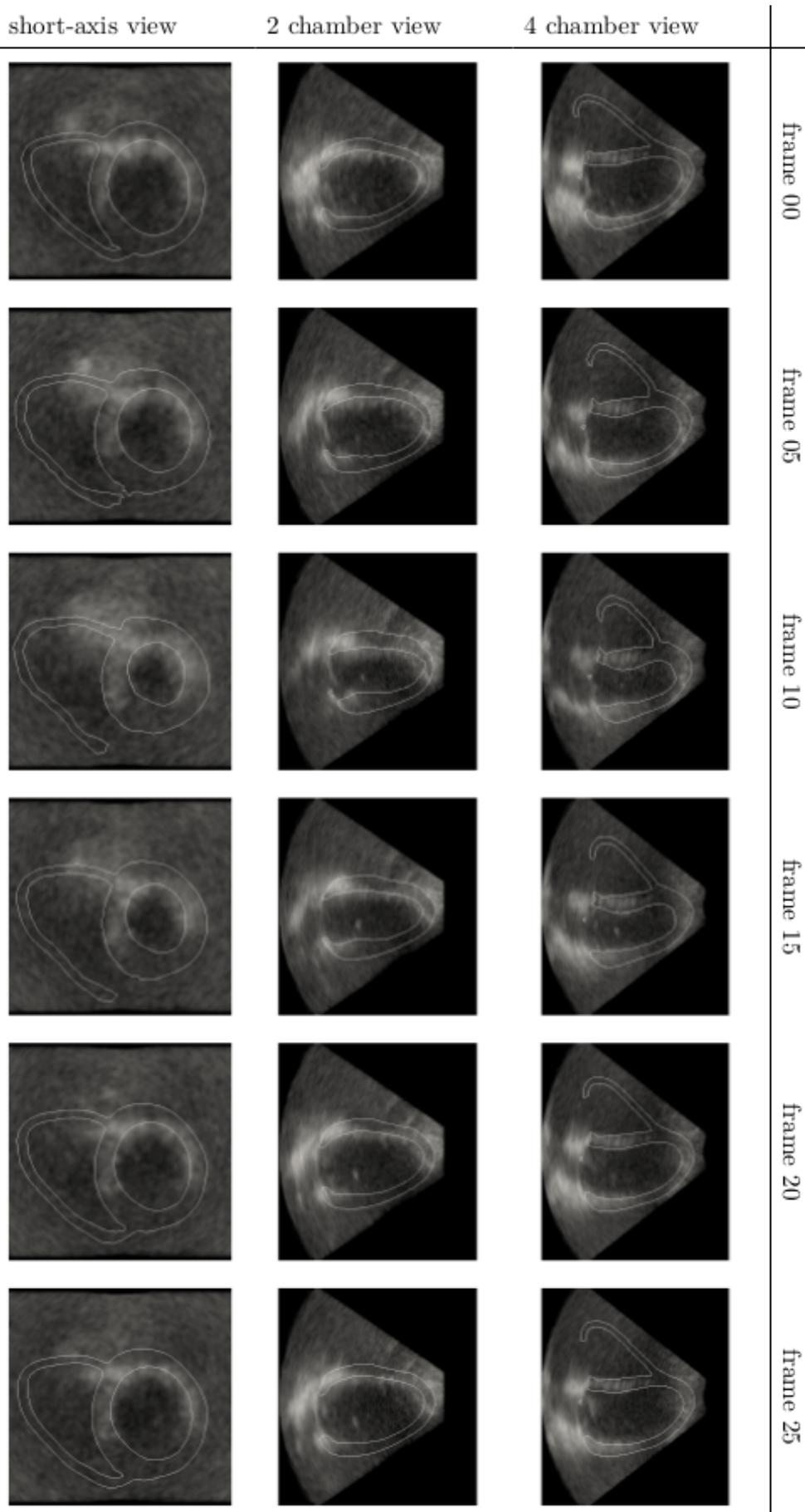
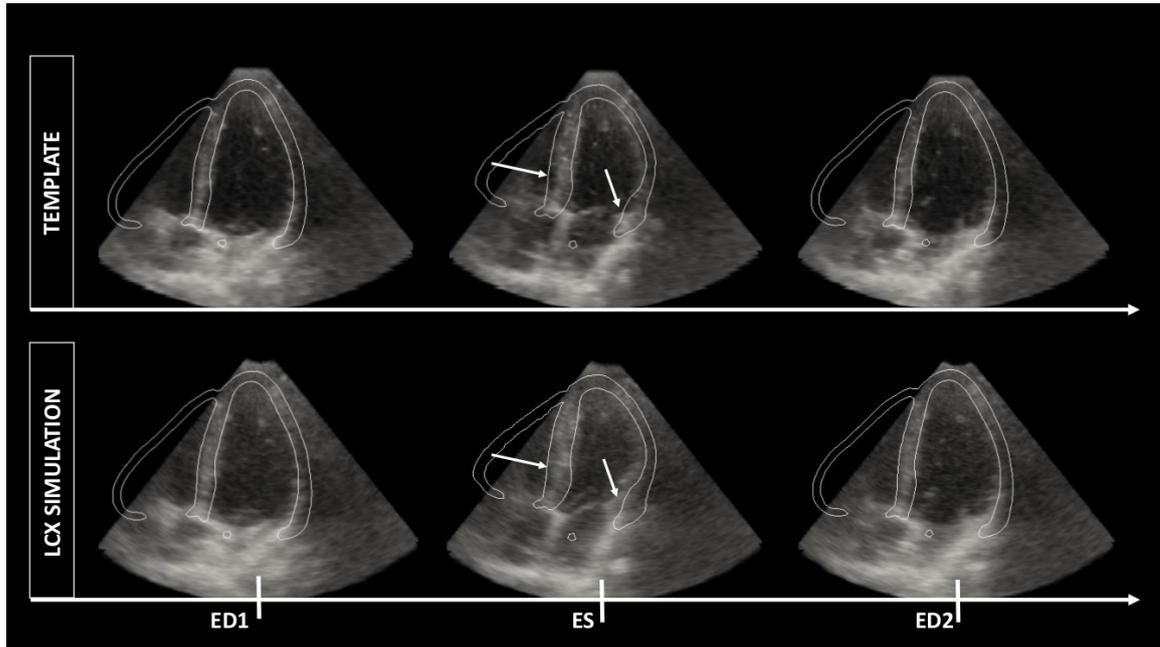
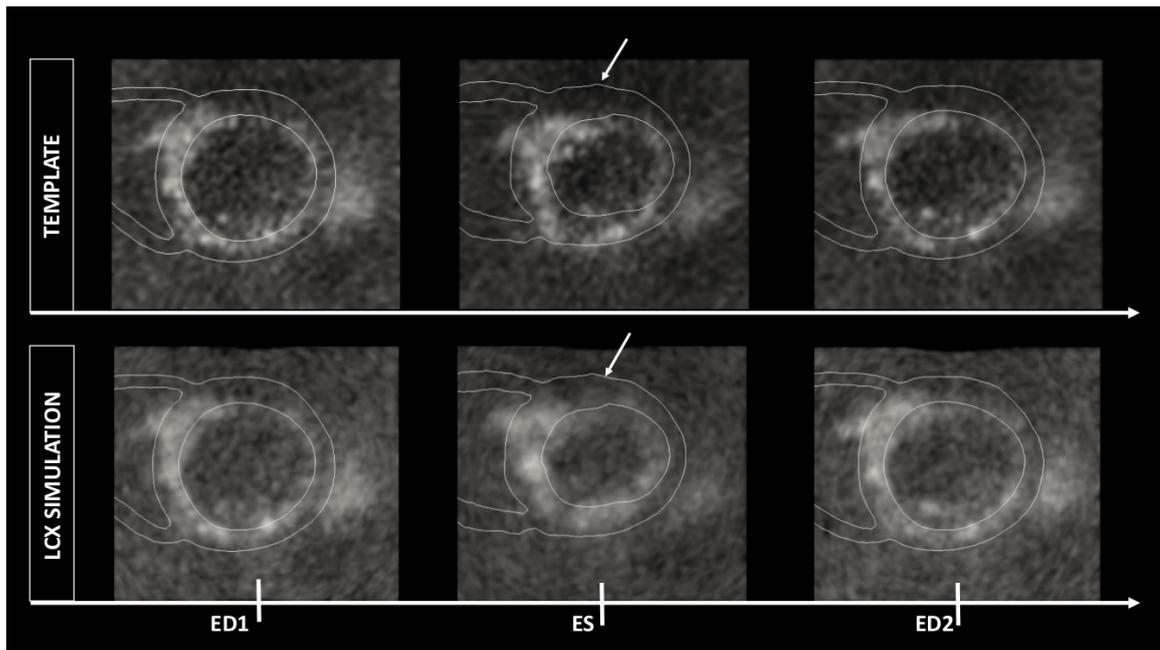


FIGURE 9 – Simulated ultrasound images. The 4-chamber, 2-chamber and short-axis views are shown. The mesh contour is superposed on the images.

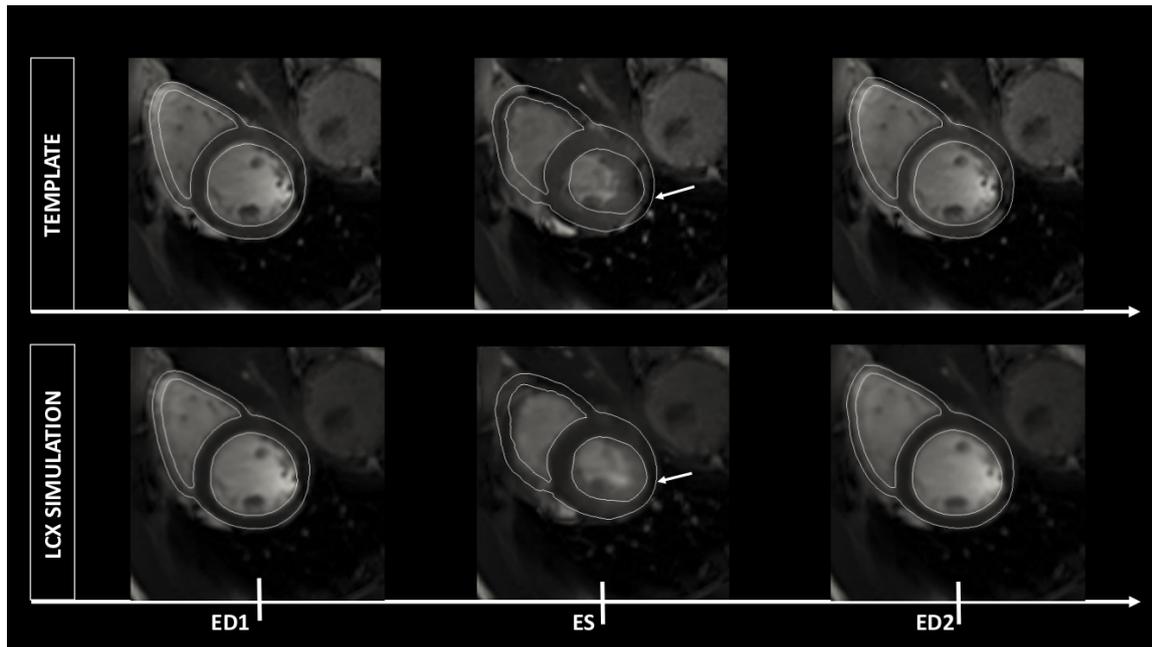


(a) four chamber view, US

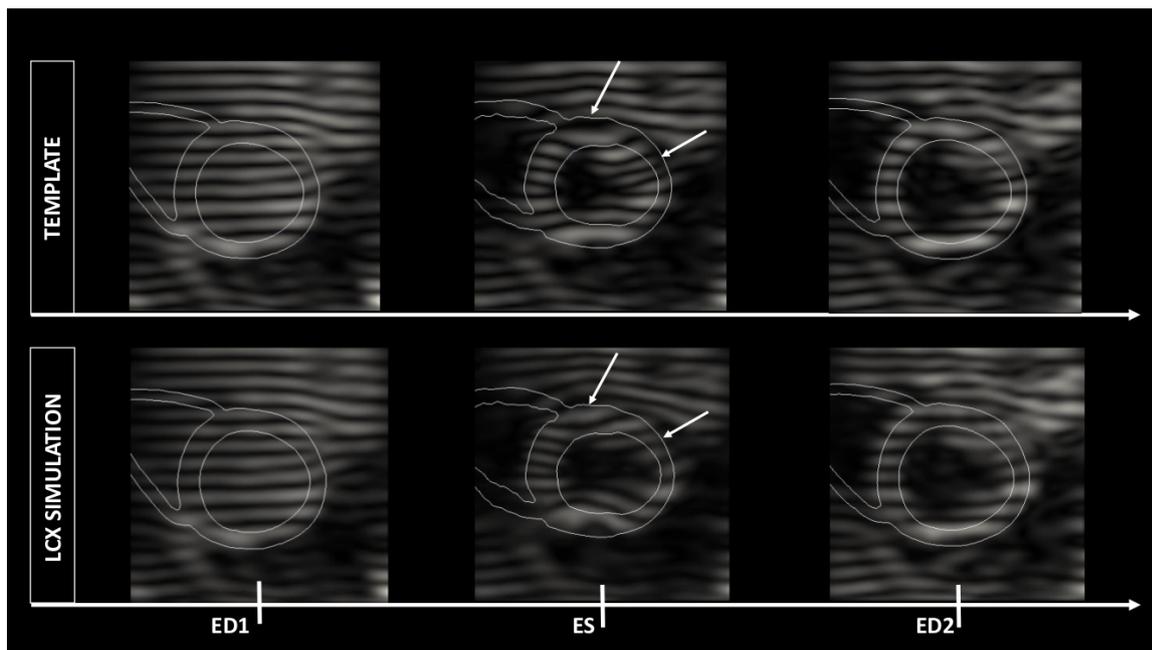


(b) short-axis view, US

FIGURE 10 – Template US recording (top row) and LCX simulation (bottom row) at matched time instants for an apical four chamber (a) and a short axis (b) slice. The E/M meshes (i.e. ground truth segmentation of the LCX simulation) are superimposed to both sequences as white contours. Template, simulation and, therefore, E/M geometry are aligned by construction at time zero (i.e. at ED1). From then on, the simulation runs independently of the template recording and, therefore, the E/M geometry no longer provides a valid segmentation of the latter. In the figure, this is particularly evident at regions marked by the white arrows. This is a consequence of the fact that, as due to the simulated ischemia, contraction in the E/M simulation is reduced as compared to the template.



(a) short-axis view, Cine MR



(b) short-axis view, Tagged MR

FIGURE 11 – Template MR recording (top row) and LCX simulation (bottom row) at matched time instants for a short-axis cine MR slice (a) and a short-axis tagged MR slice (b). The E/M meshes (i.e. ground truth segmentation of the LCX simulation) are superimposed to both sequences as white contours. Template, simulation and, therefore, E/M geometry are aligned by construction at time zero (i.e. at ED1). From then on, the simulation runs independently of the template recording and, therefore, the E/M geometry no longer provides a valid segmentation of the latter. In the figure, this is particularly evident at regions marked by the white arrows. This is a consequence of the fact that, as due to the simulated ischemia, contraction in the E/M simulation is reduced as compared to the template.

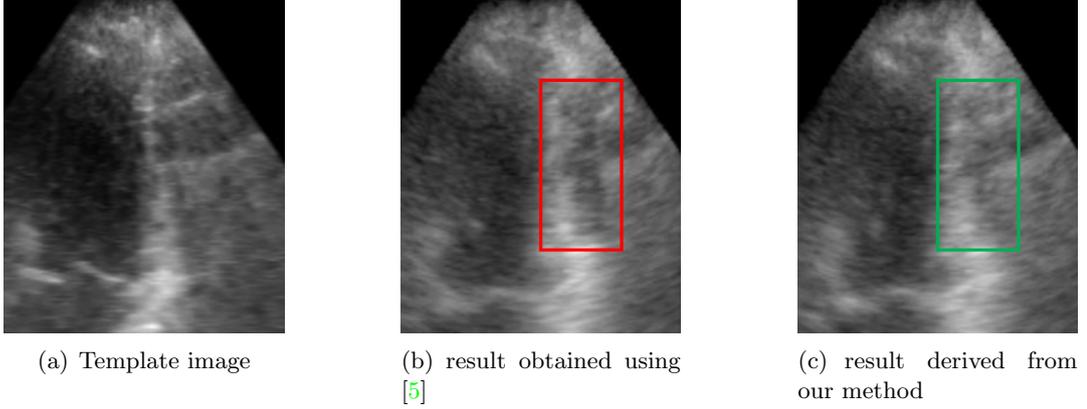


FIGURE 12 – Illustration of the interest of the proposed method to better handle the interface between myocardium and the surrounding structures. a) template image at ES; b) simulated image using [5]; c) simulated image using our approach.

T2 and proton density maps were directly computed from the corresponding template frame. This result is further confirmed by two statistical tests. Both the Wilcoxon signed-rank test and the Levene’s test returned p-value greater than 0.05, meaning that the median and variance of the myocardial intensities are statistically similar. At ES, the histograms are still in a good agreement, even if some slight difference may be observed. We applied the same statistic tests on the myocardial intensities. The Wilcoxon signed-rank test returned a p-value greater than 0.05, revealing that the template and the simulation have similar median values. However, the Levene’s test returned a p-value below 0.05, meaning that the variances are statistically different. This is coherent with what we see from the histograms. This is due to the fact that intensities of the T1 and T2 maps were not updated over the cardiac cycle, only warping strategies were performed. Moreover, the myocardial intensity changes were embedded through the proton density maps which was modeled through a simple linear fit. Nevertheless, the similarity of the distributions validates the proposed strategy to handle the myocardial intensity changes while preserving the motion derived from the E/M model. The myocardial histograms of tagged MR simulations are also shown in Fig. 13. Results show that the exact same observations can also be made for this modality.

4.8.3 US image properties : quantitative assessment

Speckle statistics : The level of realism of the simulated US images was assessed quantitatively by measuring the values of image features commonly exploited for the analysis of cardiac ultrasound datasets. Apart from pixel intensity, first order speckle statistics represent a reliable feature in common tasks as segmentation [31], [6] and motion estimation [249], [173]. In particular, while blood returns *fully developed speckle*, leading to the popular Rayleigh model [277], myocardial tissue generates instead *partially developed speckle*, leading to sub-Rayleigh distributions [34], [31], where the main factor influencing the type of statistical distribution is the number of scattering sites per resolution cell of the ultrasound system [277]. Hereto the local concentration of scatterers was optimized in order to obtain a realistic statistical description of the simulated speckle patterns. Experimentally we verified that a spatially uniform and sufficiently high concentration of scatterers (*i.e.* > 4 scatterers per resolution cell) was sufficient to generate realistic speckle statistics. Indeed, since scattering amplitude is directly sampled

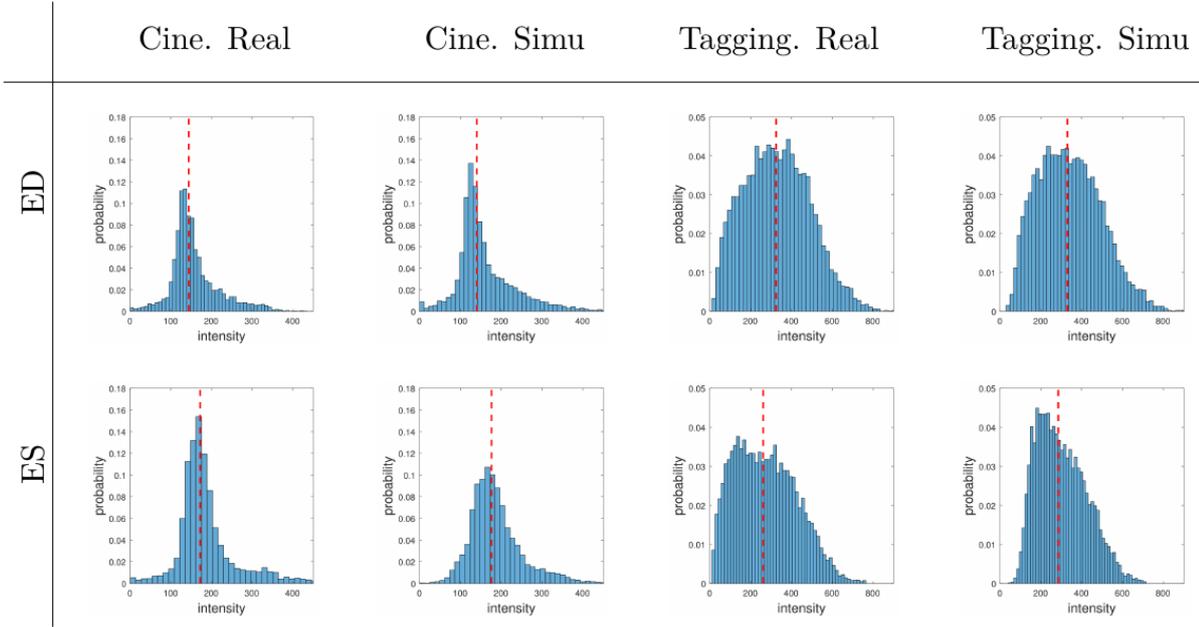


FIGURE 13 – Histograms of myocardial intensities in both real and simulated MR data. Red dashed lines mark median intensities. Cine MR results are shown on the left side and tagged MR shown on the right side.

from a speckle image (*i.e.* the template frame), the desired statistical properties are directly inherited from the latter provided the sampling is dense enough. The local histograms for two image patches extracted from the blood pool and the myocardium are reported in Fig. 14. Each analyzed ROI was taken equal to three times the size of the resolution cell in each direction and comprised ~ 6800 samples. Note that histograms were computed on data before subsampling and scan conversion (hence the high number of samples). The analysis showed that computed histograms are in good agreement with the best-fitting distributions known from the literature.

Speckle decorrelation : Speckle patterns decorrelate over time due to tissue motion and deformation. This is mainly due to the changes in the relative distance between the scattering centers (*i.e.* the strain) and, in particular in the case of sectorial probes of use in echocardiography, to the spatial variations of the system PSF [262]. Motion decorrelation weakens the fundamental assumption of many speckle tracking solutions and might negatively impact their performance. Hereto, the level of motion decorrelation present in the proposed sequences was measured. Note that motion decorrelation is naturally ensured by the proposed framework thanks to the decoupling of the mechanical simulation, establishing the motion of the individual scatterers, from the ultrasound simulation. This is a further advantage from the work in [205] where, by applying the deformation directly to the intensity images, speckle correlation was unrealistically enforced.

We evaluated motion decorrelation by considering eight 3D blocks distributed uniformly along the LV. Blocks were centered on an apical four chamber slice of the simulated 3D volume and their size was $8 \times 4 \times 4 \text{ mm}^3$ ($14 \times 6 \times 8$ voxels). The pixels inside each block were then propagated independently over time by using the reference displacement field (*i.e.* the real tissue motion from the simulated E/M model). The correlation coefficient ρ between the deformed

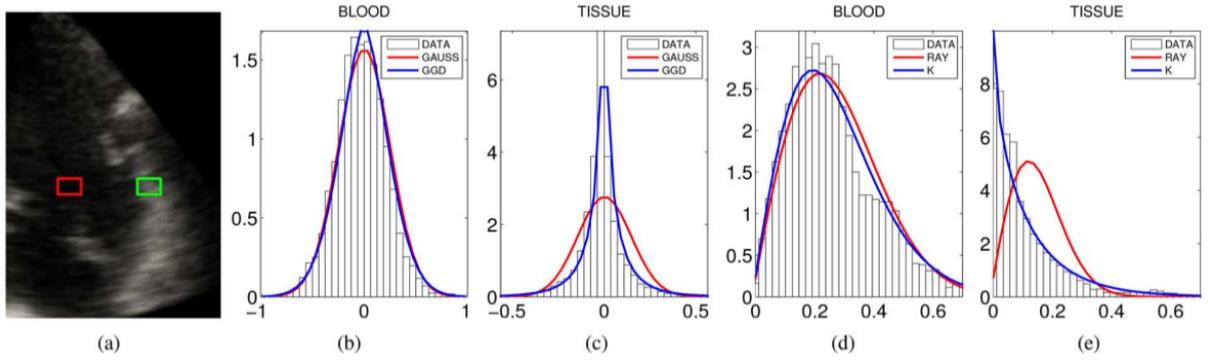


FIGURE 14 – Statistical distribution for two rectangular speckle patches belonging to the blood pool and the myocardium (red and green blocks in (a)). Each ROI is rectangular with side equal to three times the size of the system PSF in the corresponding direction and comprises ~ 6800 samples. Note that, although voxel data is reported in (a) for a more intuitive visualization, histograms were computed on raw data, *i.e.* before down-sampling and scan conversion. In (b) and (c) the histograms of the RF signal. In (d) and (e) the histograms for the non-compressed envelope signal. In the fitting of the RF and envelope histograms Gaussian and Rayleigh distributions show good fit respectively for fully developed speckle regions (*i.e.* the blood pool) while generalized Gaussian distribution (GGD) [34] and K-distribution [31] were adopted for sub-Rayleigh regions (*i.e.* the tissue). Consistently with what is expected on real recordings, Gaussian and Rayleigh models correctly fit the amplitude data inside the blood pool while speckle inside the muscle follows instead more heavily tailed distributions. Note that histograms have been normalized to have unitary area in order to be comparable with the associated probability density functions.

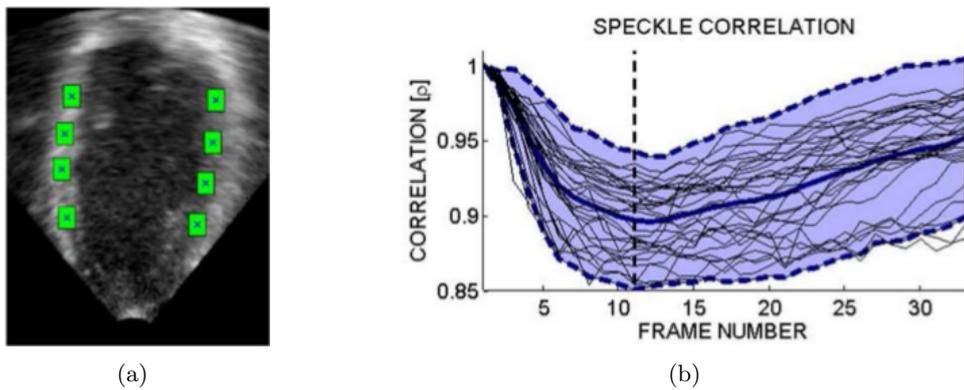


FIGURE 15 – a) Eight 3D patches used to evaluate speckle decorrelation. b) Time evolution of the correlation coefficient : in black the measurements within each block, in blue the confidence interval ($\mu \pm \sigma$). The dashed vertical line denotes block, in blue the confidence interval end of systole.

block and the initial one was then measured. The experiment was repeated for the healthy and the ischemic sequences. The obtained correlation curves are reported along with the confidence interval in Fig. 15. As expected in a real setting, correlation decreases until end of systole, when both the change in the distance between the scatterers (as due to strain) and the displacement *w.r.t.* the original position (hence, the PSF variation) are the largest. These results are qualitatively in agreement with previous simulation studies [262]. Consider that [205] would give by construction correlation values constantly equal to one (absent interpolation artifacts).

4.8.4 Mechanical properties

Global indices : Table 2 displays the LV volumes at ED and ES along with the corresponding EF for the healthy and pathological cases. Those values were computed from the simulated meshes \mathcal{S}_k used as benchmark for the proposed database. For the healthy case, the LV volumes present values around 104 mL at ED and 50 mL at ES, which is in agreement with the reference range published in [202] when combining all ethnic groups, all ages and both sexes in the analysis. In the case of dyssynchrony, the ES volumes slightly increase to values around 53 mL, revealing a reduced blood output. This tendency is clearer for ischemic cases where ES volumes are above 60 mL. Regarding the EF, the values are around 52% for healthy cases and concentrated around 40% for ischemic cases. The EF of the healthy cases are rather close to but slightly below the reference values from the literature (above 60%) [12].

TABLE 2 – Average LV volumes at ED/ES and the corresponding EF distributed among the virtual population.

	Healthy	LBBB	LCX	LAD	RCA	RCA2
ED (mL)	104.0	104.0	104.0	104.0	104.0	104.0
ES (mL)	49.6	52.8	65.7	60.9	60.3	63.4
EF (%)	52.3	49.2	36.8	41.4	42.0	39.0

Strain Curves : We show in Fig. 16 the full set of regional strain (radial, longitudinal and circumferential) curves at the basal, mid and apical levels of the LV for one healthy and four pathological cases. The regional strain curves were computed using engineering strain as in [5]. Their properties are in accordance with the ones described in the literature. For instance, in the LBBB case, the ES timings differ between the septum (AHA n° 2,3,8,9,14) and the lateral wall (AHA n° 5,6,11,12,16). The septum is activated slightly before the lateral wall, meaning the myocardium contracts asynchronously. Besides, ischemic cases show wide disparities between healthy and pathological segments. Reduced deformation values are observed for those infarcted AHA segments. In addition, we show in Fig. 17 the global circumferential strain (GCS) for the 18 virtual patients simulated in this paper. We observe that pathological cases display reduced circumferential deformations compared to healthy ones.

4. 1st contribution : realistic synthetic US and MRI sequences from the same virtual patients

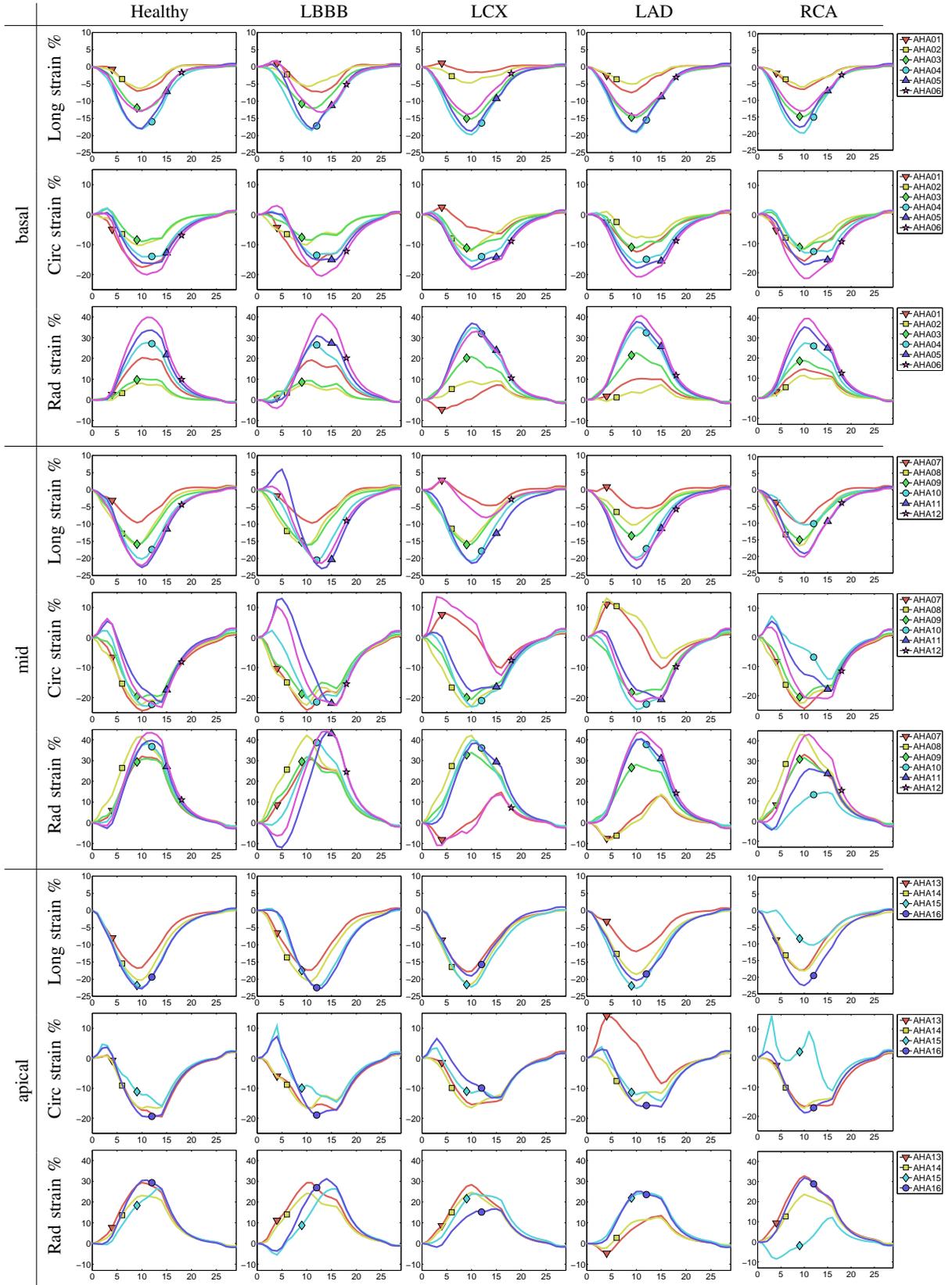


FIGURE 16 – Regional strain curves at basal, mid and apical levels of the LV. We show one healthy, one dyssynchrony (LBBB) and three infarcted hearts (LCX, LAD, RCA).

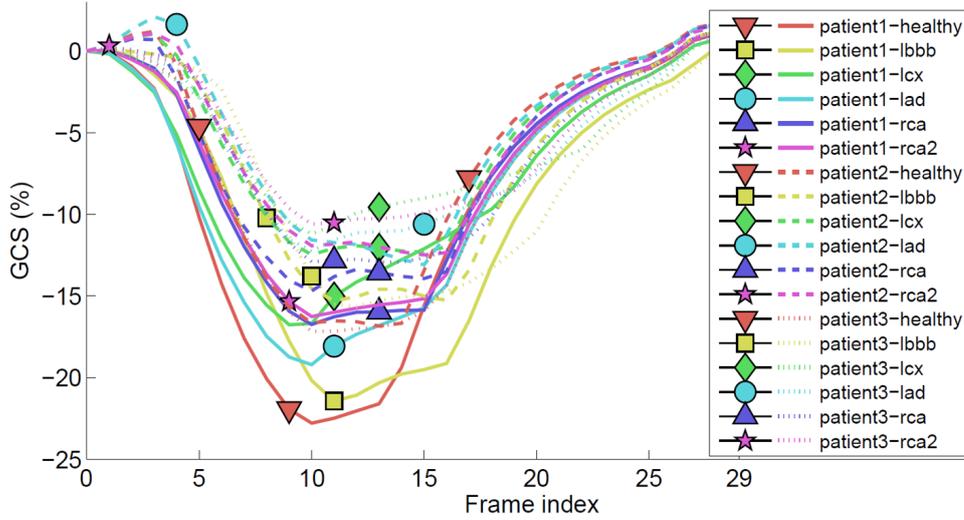


FIGURE 17 – Global circumferential strain curves corresponding to the 18 simulated virtual patients.

4.9 Discussion

It is widely acknowledged that multimodality raises many issues in cardiac imaging, beyond the mapping of all information in a common reference space. Indeed, different modalities were reported to provide significant differences for the quantification of volumes and ejection fraction values [114]. Offering to the community a consolidated dataset, on the basis of which generic or modality-specific algorithms can be quantitatively compared, is a step forward towards understanding whether reported differences between modalities originate from the processing, the difference in image quality or can only be explained by more fundamental differences in the physics of imaging the same complex tissue though different modalities. Our simulation method ensures that the synthetic MR and US sequences are generated from the same patient. The E/M simulations generated in the MR system of coordinates were brought to the US system by a simple rigid transformation composed of translations and rotations. The interest of such rigid transformation resides in the fact that it will not introduce any local myocardium deformation. The deformation patterns represented in the US and MR sequences are therefore the same, allowing a straightforward cross-modality comparison. However, this rigid transformation is sensitive to the accuracy of the 24 manual landmarks (see Sec. 4.3). The manual landmarking should thus be carefully performed in order to minimize the potential impact on the US simulations.

By combining warpings of the whole template sequence and of the first frame only to each simulation mesh, we obtained a realistic background while guaranteeing that no registration error can introduce artifacts inside the myocardium. This is a fundamental difference with all registration-based methods such as [205] where the error on displacements gets composed with the true displacement field, making any derived benchmark affected by such artifacts inside the myocardium. Moreover, we introduced an efficient Gaussian weighting function to deal with the interface between the myocardium and the background. Such a scheme guarantees a smooth transition between the different structures around the myocardium and thus avoids potential artifacts as it may appear for instance from the pipeline developed in [5] (Fig. 12).

Despite the advantages, we also acknowledge a number of limitations and simplifications.

First, the realism of the computed relaxation times T_1 and T_2 could be improved to account for the temporal variations of these quantities as reported in [220]. The corresponding spatial distribution we used was kept simple (average reference values + Gaussian noise for the background region). Also, the myocardial T_1 values we computed from real tagging templates should be compared against real T_1 mapping acquisitions to validate this process. Additionally, when computing the proton density for tagged MR simulation, we assumed the effective spin-spin relaxation time T_2^* to equal the theoretical T_2 value. This assumption simplifies the simulation method and since they have close values, it should not significantly impact the simulation results. It should be emphasized however that our modeling strategy was aiming at optimizing the realism of the generated images rather than rigorously following the underlying physics. With this in mind, the derivations of most MR parameters we made from real MR images ensure a high level of realism of the simulated images as shown in Fig. 7, 8 and 9.

We put a lot of effort in this work to make sure that tracking errors would not introduce any artifacts inside the myocardium. However, it is worth mentioning that the quality of the simulated data remains sensitive to the tracking accuracy. Since the background is taken from a frame of the template sequence, tracking errors will appear as a residual motion applied to the background, resulting in possible tag line fractures occurring at the myocardium/background interface, although the use of a Gaussian weighting function is expected to limit such unrealistic artifacts. An accurate tracking is therefore required for producing high quality simulation data, which we ensured by using high quality images from carefully selected healthy subjects [258].

Finally, the median longitudinal strain corresponding to the cardiac deformations generated by the E/M model is around 15%, which is still lower than reference clinical values. This has already been discussed in [5] and remains one of the major limitations of the E/M model. Compared to [5], the simulations generated in this paper already show increased longitudinal deformations thanks to a thorough parameter tuning process. Further improvement of the longitudinal strain as well as the reduction of the dispersion of GCS and the increase of the EF remain the object of future work.

4.10 Conclusions

We introduced in this study a generic framework allowing the generation of realistic multimodal (US, cine and tagged MR) cardiac synthetic sequences for the same virtual patient. The following novelties were introduced in this study :

- the combination of an E/M model with US/MR physical simulators to introduce the possibility of interacting with the image formation process ;
- the use of multimodal template sequences from the same patient to extract the most relevant information from each modality ;
- the introduction of a novel combinative warping strategy, with the goal of reducing motion artifacts that may occur in myocardium regions ;
- the contribution of the first unified framework which allows the generation of multimodal (US, cine and tagged MR) realistic synthetic sequences for the same virtual patient ;
- the use of a novel strategy based on a Gaussian weighting function to generate a smooth interface between the myocardium and the surrounding structures.

5 2nd contribution : anatomical deformable model for strain estimation in cardiac imaging

In this section, we present a novel framework based on an anatomical deformable model to efficiently quantify myocardial motion field and strain at once. Myocardial motion tracking is regarded as a succession of image-to-image registration problem. To take advantage of the temporal coherence of the sequence, we chose to adopt a frame-to-frame strategy. The performance of the proposed method is assessed on the synthetic database presented in Sec. 4.

5.1 Motivations

Based on the literature review we made in Sec. 2, we developed a novel method integrating the advantages demonstrated by several available strategies. Inspired by the work proposed by Heyde *et al.* [105], we decided to express the cardiac deformations directly from the myocardial space. By doing so, the integration of any prior such as local motion or regularization will be efficiently adapted to the anatomy of the heart. One of the conceptual key of the proposed formulation is the regularization of the estimated strain and motion displacement according to a windows-weighted regression method which maintains a low computational complexity. Based on the observation that radial direction is the most challenging to recover meaningful strains in both tagged MR and US, we decided to enforce myocardial incompressibility by correcting strain values in the radial direction only. This is likely due to the fact that the number of tags from endocardium to epicardium is small in the radial direction compared to circumferential and longitudinal directions in tagged MR, while the resolution in the lateral and azimuth directions is lower than in the axial direction in ultrasound imaging. Finally, it should be emphasized that the developed approach has the advantage of being generic and can be easily adapted and applied to any modality. In the sequel, we describe the key aspects of the proposed formalism.

5.2 Anatomical mesh model

An anatomical mesh is first initialized from the 3D segmentation of the myocardium at ED using any algorithms. The segmentation output can be either a binary mask or a closed surface mesh that encompasses LV endocardium and epicardium. From this result, a resampling procedure is then applied for two main purposes : *i*) the surface mesh is converted to a volumetric dense mesh which also contains an inner layer between endocardium and epicardium ; *ii*) the LV domain is sampled regularly in radial, longitudinal and circumferential directions. This facilitates regularizing the displacement field on the mesh topology.

5.2.1 Resampling rules and directional indexes

The resampling is done by casting rays from the long axis to the endocardial and epicardial surfaces as described in Fig. 18. The mid part of the LV is discretized using a cylindrical sampling. The long axis is sampled uniformly. From each long axis point, N_c rays are then cast with equal angular binning in the plane orthogonal to the long axis. For each ray, the intersection with the surface mesh gives two points : one at endocardium and one at epicardium. The average of these two points produces the inner middle layer. Similarly, apical and basal parts are sampled in a spherical manner. Azimuth and elevation angles are binned uniformly (see Fig. 18-a and 18-b).

Following this discretization scheme, each point in the generated mesh can be associated to three indexes. First, each point is either endocardial, mid or epicardial. Second, each point is

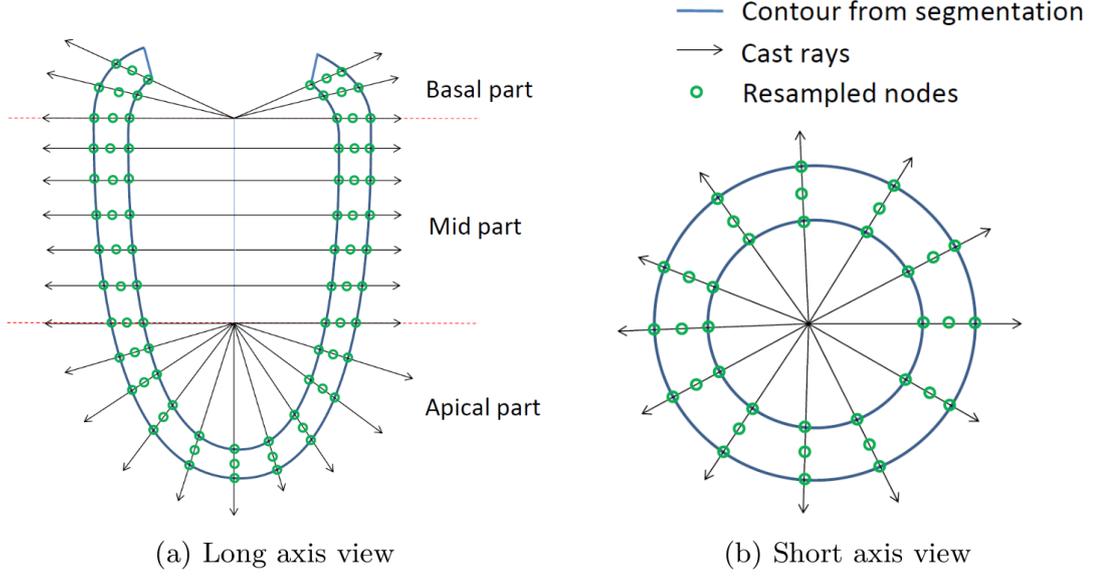


FIGURE 18 – Resampling of LV surface mesh to volumetric dense mesh.

intercepted by a single ray in the *circumferential* direction. Finally, each ray is either casted from one source point on the long axis, being indexed from the sequential cylindrical discretization or angled with respect to the long axis, being indexed from the sequential spherical discretization (Fig. 18-a). In this way, every node in the mesh is related to three indexes (m_r, m_l, m_c) where $m_d \in [1, N_d]$ and N_d ($d = r, l, c$) is the number of divisions in each of the three directions (*radial, longitudinal or circumferential*). Then the spatial coordinates of every node can be written as $\mathcal{P}(m_r, m_l, m_c)$. In our experiments, $N_r = 3$, $N_l = 35$ and $N_c = 50$. As the number of tags in tagged MR and the resolution in US are low in the radial direction, we chose to use only three radial layers. N_l and N_c were chosen to produce isotropic cells in the volumetric mesh, which implies similar cell resolutions in longitudinal and circumferential directions. By convention, radial indexes are ordered from endocardium to epicardium, longitudinal indexes are ordered from apex to base, and circumferential indexes go counter-clockwise when looking from the apex.

5.2.2 Local directions

Based on the mapping \mathcal{P} , local *radial, longitudinal* and *circumferential* directions $\hat{\mathbf{e}}_r(\mathbf{x})$, $\hat{\mathbf{e}}_l(\mathbf{x})$ and $\hat{\mathbf{e}}_c(\mathbf{x})$ can be computed by normalizing the edges of the volumetric mesh. However, this set of directions would not be orthogonal. To ensure local orthogonality directions at all locations except the purely apical ones during the regularization procedure, local directions were defined as follows. For a non-apical mesh point \mathbf{x} whose directional indexes are (m_r, m_l, m_c) , its longitudinal direction can be computed as :

$$\begin{aligned} \mathbf{e}_l(\mathbf{x}) &= \mathcal{P}(m_r, m_l, m_c) - \mathcal{P}(m_r, m_l - 1, m_c) \quad , \text{ if } m_l > 1 \\ \hat{\mathbf{e}}_l(x) &= \frac{\mathbf{e}_l(x)}{\|\mathbf{e}_l(x)\|} \end{aligned} \quad (17)$$

Its circumferential direction $\hat{\mathbf{e}}_c(\mathbf{x})$ was then computed similarly but further corrected to be

orthogonal to $\hat{\mathbf{e}}_l(\mathbf{x})$:

$$\begin{aligned} \mathbf{e}_c(\mathbf{x}) &= \mathcal{P}(m_r, m_l, \text{mod}(m_c + 1, N_c)) - \mathcal{P}(m_r, m_l, m_c) \quad , \text{ if } m_l > 1 \\ \hat{\mathbf{e}}_c(\mathbf{x}) &= \frac{\mathbf{e}_c(\mathbf{x}) - (\mathbf{e}_c(\mathbf{x}) \cdot \hat{\mathbf{e}}_l(\mathbf{x}))\hat{\mathbf{e}}_l(\mathbf{x})}{\|\mathbf{e}_c(\mathbf{x}) - (\mathbf{e}_c(\mathbf{x}) \cdot \hat{\mathbf{e}}_l(\mathbf{x}))\hat{\mathbf{e}}_l(\mathbf{x})\|} \end{aligned} \quad (18)$$

Finally, the radial direction $\hat{\mathbf{e}}_r(\mathbf{x})$ was obtained from the cross product between $\hat{\mathbf{e}}_c(\mathbf{x})$ and $\hat{\mathbf{e}}_l(\mathbf{x})$.

$$\hat{\mathbf{e}}_r(\mathbf{x}) = \hat{\mathbf{e}}_c(\mathbf{x}) \times \hat{\mathbf{e}}_l(\mathbf{x}) \quad , \text{ if } m_l > 1 \quad (19)$$

For purely apical points ($m_l = 1$), $\hat{\mathbf{e}}_c(\mathbf{x})$ has no definition due to the singularity on the apex. $\hat{\mathbf{e}}_r(\mathbf{x})$ was defined in a similar manner by normalizing the edges in the radial direction. $\hat{\mathbf{e}}_l(\mathbf{x})$ was computed by taking the mean of its N_c nearest neighbors located in the same radial layer and normalizing the result.

5.2.3 Anatomical coordinates

Our goal is to map Cartesian (\mathbf{x}) to anatomical (r, l, c) coordinates while satisfying the following properties :

$$\begin{aligned} \forall \mathbf{x} \in \Omega, \quad \nabla r(\mathbf{x}) &= \hat{\mathbf{e}}_r(\mathbf{x}), \\ \nabla l(\mathbf{x}) &= \hat{\mathbf{e}}_l(\mathbf{x}), \\ \nabla c(\mathbf{x}) &= \frac{1}{\rho(\mathbf{x})}\hat{\mathbf{e}}_c(\mathbf{x}) \quad . \end{aligned} \quad (20)$$

where Ω is the LV myocardium domain, and $\rho(\mathbf{x})$ is the distance of point \mathbf{x} to the LV long axis.

For each direction $d \in \{r, l, c\}$, we start from a group of reference points $\{X_0^d\}$ where the anatomical coordinates are set to be zero. For longitudinal coordinates, the apex is taken as a reference point. In the circumferential direction, points located at the center of the local windows (described later in Sec. 5.3.1) are chosen as reference points. Finally, for the radial direction, $\{X_0^r\}$ is defined as the set of mid-layer points.

From $\{X_0^d\}$, the coordinates are obtained by curvilinear integration. Taking one direction $d \in \{r, l, c\}$, $d(\mathbf{x})$ is defined as

$$d(\mathbf{x}) = \begin{cases} \min_{\mathcal{C}_d(\mathbf{x})} \int_{\mathcal{C}_d(\mathbf{x})} \hat{\mathbf{e}}_d(\mathbf{s}) \cdot d\mathbf{s}, & \text{if } d = r, l \\ \min_{\mathcal{C}_d(\mathbf{x})} \int_{\mathcal{C}_d(\mathbf{x})} \frac{1}{\rho(\mathbf{s})} \hat{\mathbf{e}}_d(\mathbf{s}) \cdot d\mathbf{s}, & \text{if } d = c \quad . \end{cases} \quad (21)$$

where $\mathcal{C}_d(\mathbf{x})$ is a geodesic path on the mesh starting at a node $\in X_0^d$.

The $r(\mathbf{x})$ and $l(\mathbf{x})$ coordinates are computed globally for the whole LV. However, practically, the $c^{(i)}(\mathbf{x})$ coordinate is defined locally for each window i (described later in Sec. 5.3.1) in order to ensure its continuity over the window domain. Examples of r, l over the whole myocardial domain and $c^{(i)}$ over the local window i are plotted in Fig. 19.

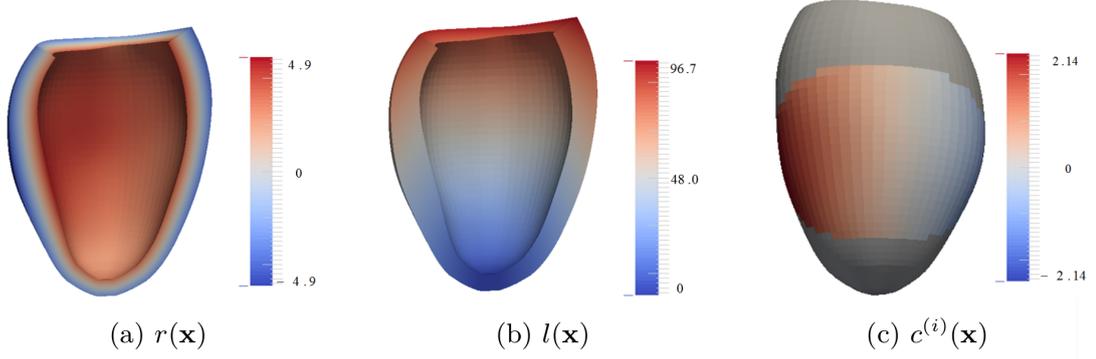


FIGURE 19 – Anatomical coordinates $r, l, c^{(i)}$ defined on the LV volumetric mesh (in (c), $c^{(i)}(\mathbf{x})$ has no definition in the surrounding grey region since it is beyond the range of the current window i).

5.3 Tracking methodology

5.3.1 Anatomical deformation model

Motion model : Let $\mathbf{u}_t^{(i)}(\mathbf{x})$ be a local parametric motion within a window i (described later in this section) from $t \rightarrow t + 1$. As tracking aims at recovering both motion and deformation values, we opted for an affine transformation in the anatomical coordinates $r, l, c^{(i)}$ that combines translations, stretching (or shortening) and shearing. This 12-parameter model was defined in the anatomical system of coordinates introduced in Sec. 5.2 :

$$\begin{aligned}
 \mathbf{u}_t^{(i)}(\mathbf{x}) &= L_r^{(i)}(\mathbf{x})\hat{\mathbf{e}}_r(\mathbf{x}) + L_l^{(i)}(\mathbf{x})\hat{\mathbf{e}}_l(\mathbf{x}) + L_c^{(i)}(\mathbf{x})\hat{\mathbf{e}}_c(\mathbf{x}) \\
 \text{with } L_r^{(i)}(\mathbf{x}) &= a_{rr}^{(i)}r(\mathbf{x}) + a_{rl}^{(i)}l(\mathbf{x}) + a_{rc}^{(i)}c^{(i)}(\mathbf{x}) + b_r^{(i)} \\
 L_l^{(i)}(\mathbf{x}) &= a_{lr}^{(i)}r(\mathbf{x}) + a_{ll}^{(i)}l(\mathbf{x}) + a_{lc}^{(i)}c^{(i)}(\mathbf{x}) + b_l^{(i)} \\
 L_c^{(i)}(\mathbf{x}) &= (a_{cr}^{(i)}r(\mathbf{x}) + a_{cl}^{(i)}l(\mathbf{x}) + a_{cc}^{(i)}c^{(i)}(\mathbf{x}) + b_c^{(i)})\rho(\mathbf{x})
 \end{aligned} \tag{22}$$

The set of zero-order $\{b_r^{(i)}, b_l^{(i)}, b_c^{(i)}\}$ (related to motion), first-order $\{a_{rr}^{(i)}, a_{ll}^{(i)}, a_{cc}^{(i)}\}$ (related to *radial, longitudinal* and *circumferential* strains) and first-order $\{a_{rl}^{(i)}, a_{rc}^{(i)}, a_{lr}^{(i)}, a_{lc}^{(i)}, a_{cr}^{(i)}, a_{cl}^{(i)}\}$ (related to shearing) parameters fully define the motion in the local window i .

The partition of unity method : As introduced in [76], a global regression problem can be simplified by solving a set of local ones with the help of the partition of unity technique. A partition of the LV domain is defined by a set of control points $\{\mathbf{x}_i\}$ and by window functions $\varphi^{(i)}(\mathbf{x})$ associated to each control point. The latter are assumed to satisfy the following conditions :

$$\begin{aligned}
 \arg \max_{\mathbf{x}} \varphi^{(i)}(\mathbf{x}) &= \mathbf{x}_i, \quad \forall i \\
 \varphi^{(i)}(\mathbf{x}) &\geq 0, \quad \forall \mathbf{x} \in \Omega, \\
 \sum_i \varphi^{(i)}(\mathbf{x}) &= 1 \quad \forall \mathbf{x} \in \Omega.
 \end{aligned} \tag{23}$$

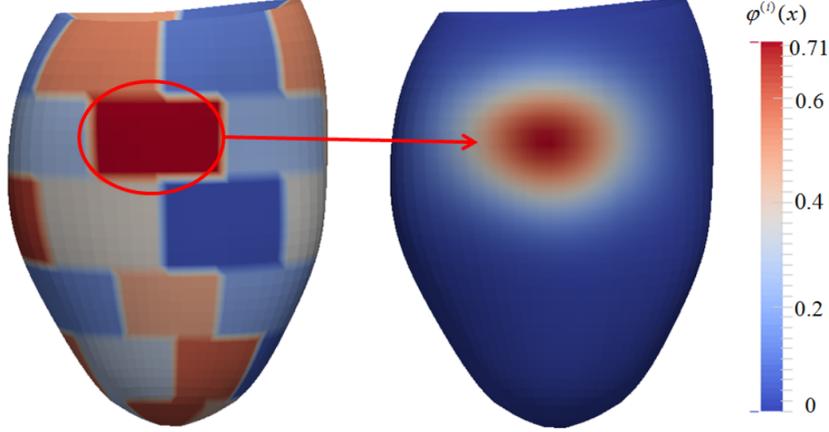


FIGURE 20 – Definition of LV local windows and an example of the window function for the red-circled window.

where Ω is the LV myocardium. The regularized displacement $\mathbf{u}_t(\mathbf{x})$ is computed as

$$\mathbf{u}_t(\mathbf{x}) = \sum_i \varphi^i(\mathbf{x}) \mathbf{u}_t^{(i)}(\mathbf{x}) \quad (24)$$

where $\mathbf{u}_t^{(i)}(\mathbf{x})$ is the motion model introduced at the beginning of this section.

Window functions : The LV myocardium domain is partitioned into a set of windows as illustrated in Fig. 20. Window functions are centered at each control point and must decrease when the distance to the control point increases. Since we opted for tracking a volumetric mesh and defined a system of coordinates on the mesh, the geodesic distance was chosen rather than the Cartesian one for computing the window function. Let us first consider a standard Gaussian kernel involving the geodesic distance ξ between \mathbf{x} and the center of the window \mathbf{x}_i .

$$g^{(i)}(\mathbf{x}) = \frac{1}{2\pi\sigma_i} \exp\left(-\frac{\xi(\mathbf{x}, \mathbf{x}_i)^2}{2\sigma_i^2}\right) \quad (25)$$

To use a single σ parameter for various window sizes, we expressed σ_i as

$$\sigma_i = \sigma\chi_i \quad (26)$$

where χ_i is computed for each window by averaging its circumferential and longitudinal extents. In Eq. (26), σ is an auxiliary parameter named as kernel bandpass. It controls how neighboring windows will overlap for producing the final displacement estimate in Eq. (24). Each window function are then defined by normalizing $g^{(i)}$ values for summing to 1 :

$$\varphi^{(i)}(\mathbf{x}) = \frac{g^{(i)}(\mathbf{x})}{\sum_j g^{(j)}(\mathbf{x})} \quad (27)$$

Myocardium incompressibility constraint : Imposing a zero divergence is a classic way of enforcing incompressibility. This constraint can be imposed either as a soft (global) or hard (local) level. In our case, computing the divergence locally would require complex spatial derivations of the $r(\mathbf{x})$, $l(\mathbf{x})$ and $c^{(i)}(\mathbf{x})$ functions. Therefore, we preferred to impose incompressibility

within each window, enforcing the overall volume change to be zero. For doing so, we applied the divergence theorem. If $\partial S^{(i)}$ stands for the surface enclosing the window i , the outward flow of $\mathbf{u}_t^{(i)}(\mathbf{x})$ through $\partial S^{(i)}$ equals the volume integral of the divergence over that window. Therefore imposing myocardial incompressibility is equivalent to forcing the flow through $\partial S^{(i)}$ to be zero. For each local window i defined on the volumetric mesh, $\partial S^{(i)}$ consists of quadrilaterals. Each of the quadrilateral j can be represented by its center \mathbf{x}_j , its normal (pointing outward) \mathbf{n}_j and its area A_j . The flow of $\mathbf{u}_t^{(i)}(\mathbf{x})$ through $\partial S^{(i)}$ can then be written as

$$\begin{aligned} Q_t^{(i)} &= \sum_{j \in \partial S^{(i)}} \mathbf{u}_t^{(i)}(\mathbf{x}_j) \cdot A_j \vec{n}_j \\ &= \sum_{j \in \partial S^{(i)}} \sum_{d=r,l,c} L_d^{(i)}(\mathbf{x}_j) (\hat{\mathbf{e}}_d(\mathbf{x}_j) \cdot A_j \vec{n}_j) \end{aligned} \quad (28)$$

Because $L_d^{(i)}(\mathbf{x}_j)$ is linear in $\{b_d^{(i)}, a_{rd}^{(i)}, a_{ld}^{(i)}, a_{cd}^{(i)}\}$ ($d = r, l, c$), so is $Q_t^{(i)}$. Imposing $Q_t^{(i)} = 0$ amounts to add a linear equality constraint. In other words, a_{rr} can be written as a linear combination of all other parameters $\{b_r^{(i)}, b_l^{(i)}, b_c^{(i)}, a_{rl}^{(i)}, a_{rc}^{(i)}, a_{lr}^{(i)}, a_{ll}^{(i)}, a_{lc}^{(i)}, a_{cr}^{(i)}, a_{cl}^{(i)}, a_{cc}^{(i)}\}$, leading to a 11-parameter motion model per window $\varphi^{(i)}$.

The final motion from $t \rightarrow t + 1$ is obtained by mixing all local motions weighted by the window functions thanks to Eq. (24). This gives a motion model with $11 \times \mathcal{N}$ parameters (\mathcal{N} being the window number). Estimating these parameters allows obtaining both anatomically coherent myocardial motions (*i.e.* locally affine) and meaningful deformations

5.3.2 Image-based registration

For the motion tracking from $t \rightarrow t + 1$, given an initial estimate of the displacement $\mathbf{u}(\mathbf{x})$, the key idea is to find an optimal incremental displacement field $\mathbf{v}(\mathbf{x}) = \sum_i \varphi^{(i)}(\mathbf{x}) \mathbf{v}^{(i)}(\mathbf{x})$ by minimizing an energy E defined as follows :

$$E(\mathbf{v}; \mathbf{u}) = \int_{\Omega} \sum_{k=0}^{K-1} \omega_k(\mathbf{x}) (\mathcal{I}_k^t(\mathbf{x}) - \mathcal{I}_k^{t+1}(\mathbf{x} + \mathbf{u}(\mathbf{x}) + \mathbf{v}(\mathbf{x})))^2 d\mathbf{x} \quad (29)$$

In the case of ultrasound or cine MR imaging : *i*) $K = 1$; *ii*) $\mathcal{I}_k^t(\mathbf{x})$ represents the image intensity at point \mathbf{x} and time t ; *iii*) $\omega_k(\mathbf{x})$ is set as a constant map equals to 1 for all pixel positions.

In the case of tagged MR imaging : *i*) $K = 3$; *ii*) $\mathcal{I}_k^t(\mathbf{x})$ represents the phase intensity at point \mathbf{x} and time t ; *iii*) $\omega_k(\mathbf{x})$ is compute from the magnitude information at point \mathbf{x} ($\mathcal{M}_k^t(\mathbf{x})$) as :

$$\omega_k(\mathbf{x}) = \sqrt{\mathcal{M}_k^t(\mathbf{x}) \mathcal{M}_k^{t+1}(\mathbf{x} + \mathbf{u}(\mathbf{x}) + \mathbf{v}(\mathbf{x}))} \quad (30)$$

5.3.3 Multi-scale framework

For capturing large myocardial motion, a multi-scale framework is used. We adopt three scales with gradually increased window numbers (*i.e.* smaller windows). The motion field is

refined each time passing from the current scale to the next. Within each scale, the procedure described above is iterated until convergence.

Algorithm 1 : Proposed multi-scale algorithm

```

Set  $n = 0$ ,  $\mathbf{u}_0 = 0$ 
for each scale  $s$  do
    Compute  $\varphi^{(i)}$ 
    repeat
        for each window  $i$  do
             $\mathbf{v}_n^{(i)} = \operatorname{argmin}_{\mathbf{v}^{(i)}} E^{(i)}(\mathbf{v}^{(i)}; \mathbf{u}_n)$ 
        end
         $\mathbf{v}_n = \sum_i \varphi^{(i)} \mathbf{v}_n^{(i)}$ 
         $\mathbf{u}_{n+1} = \mathbf{u}_n + \mathbf{v}_n$ 
         $n = n + 1$ 
    until  $\|\mathbf{v}_n\| < \delta$ ;
end

```

5.4 Results

We introduced in the previous section the computation of the motion field from $t \rightarrow t + 1$. Based on this formalism, we decided to implement a frame-to-frame tracking strategy. The motion and strain estimation of the myocardium over the whole cardiac cycle is thus regarded as a succession of frame-to-frame registration problems which are solved by the algorithm described above. Thanks to the synthetic data generated in Sec. 4, we evaluated the performance of the proposed anatomical deformable model (referred as ADM in the sequel) across the three cardiac imaging modalities (US, cine and tagged MR). This would also help to better understand the properties of each modality, especially the associated ability in quantifying local myocardial deformations. For comparison purpose, the generic motion tracking based on sparse demons algorithm was also used to process the same synthetic database [247].

For each virtual patient, we resampled the LV of the first simulation mesh (taken from the MR space) in the *radial*, *longitudinal* and *circumferential* directions by the method introduced in Sec. 5.2. The resampled mesh has three layers in the radial direction : endocardium, epicardium and a mid-myocardium layer. Special efforts were made to ensure that the resampled mesh remains inside the original simulation mesh. We show in Fig. 21 an example that illustrates the resampling results. Since the mesh was set in the MR space, as mentioned in Sec. 4.3, a rigid transformation built upon 24 manually chosen landmarks was then used to bring the resampled mesh to the US coordinate system. For each imaging modality (US, cine or tagged MR), the corresponding mesh was then tracked by both the ADM and the sparse demons algorithms. Since the sparse demons algorithm cannot handle multi-channel data, the three channels of the tagged MR image were multiplied to create a single sequence of tagging grids. To obtain the ground-truth motion of the resampled mesh points, the mesh was warped over the whole sequence by forces computed from the simulation meshes.

In total, 18 virtual patients were tracked, yielding 54 tracked mesh sequences generated by each algorithm. Similar to [5] and [299], we computed strain using the engineering approach since this metric describes the myocardial stretch/shortening in the anatomical directions that

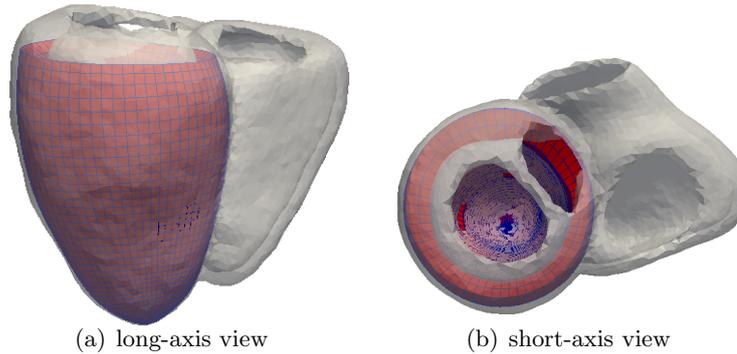


FIGURE 21 – The resampling of the first simulation mesh into a new LV mesh defined following the radial, longitudinal and circumferential directions. The resampled mesh is shown in red color. The simulation mesh is shown in white with adjusted opacity level.

are clinically relevant. In particular, for each cardiac phase, local motion and strains (radial, longitudinal and circumferential) errors were computed for each 17 AHA segments and then averaged from all time instants, all segments and all virtual patients so to produce the boxplots provided in Fig. 22. From this figure, one can first observe that the quantification of motion and strains from tagged MR and US shows smaller errors compared to cine MR. For instance, median motion errors of the ADM and the sparse demons methods are around 1 *mm* for both tagged MR and US, and increase up to 2 *mm* for cine MR. This validates the fact that motion is harder to estimate from cine MR imaging since this modality does not carry any marker in the myocardium region as it is the case in tagged MR and US imaging. From this figure, one can also observe that the two tested methods reached almost similar accuracy for the estimation of motion and longitudinal/circumferential strain on tagged MR. However, the proposed ADM model provides better results for the quantification of the radial strain on this modality. Concerning cine MR imaging, the sparse demons algorithm clearly outperformed the ADM method on both motion and strains estimation. For the US modality, the sparse demons algorithm gave more accurate motion, longitudinal/circumferential strain estimation while the ADM method performed better on radial strain thanks to the imposed radial regularization.

To study the influence of the chosen data attachment term involved in our formalism, we replaced the image intensity term in Eq. (29) by the phase information computed from the same band-pass filter as the one proposed in the HARP formalism [186]. To this end, we re-launched the two implementations (intensity and phase-based versions) with the exact same parameters and configurations on the 18 synthetic tagged MR datasets. The corresponding results are provided in Fig. 23. From the obtained boxplots, we can see that the two implementations produced very close results in terms of both motion estimation and strain quantification. This illustrates the robustness of our approach with respect to the chosen data attachment term. Indeed, by the time the data attachment term involved in Eq. (29) is properly chosen (*e.g.* either image intensity or phase information in tagged MR), the proposed ADM method will produce similar results.

5.5 Discussion

From the obtained results, ADM appears to be the most interesting tested method for quantifying tagged MR, while the sparse demons algorithm seems to be better suited for cine

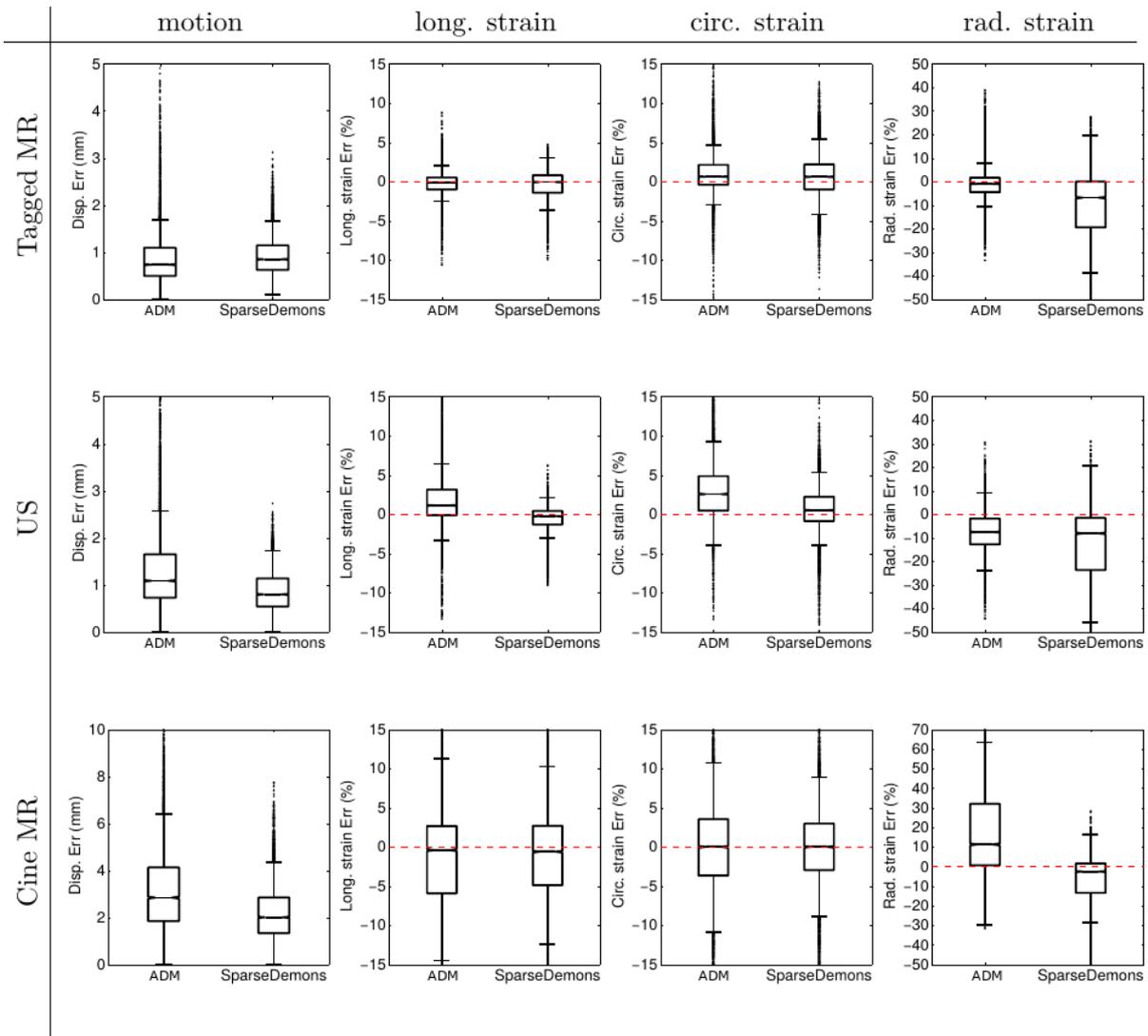


FIGURE 22 – Local motion and strain quantification errors obtained by the sparse demons and the proposed algorithms (ADM). Both methods were run on the 18 virtual patients simulated in Sec. 4. For each cardiac phase, local motion and strains errors were computed for each 17 AHA segments and then averaged from all time instants, all segments and all virtual patients so to produce the given boxplots.

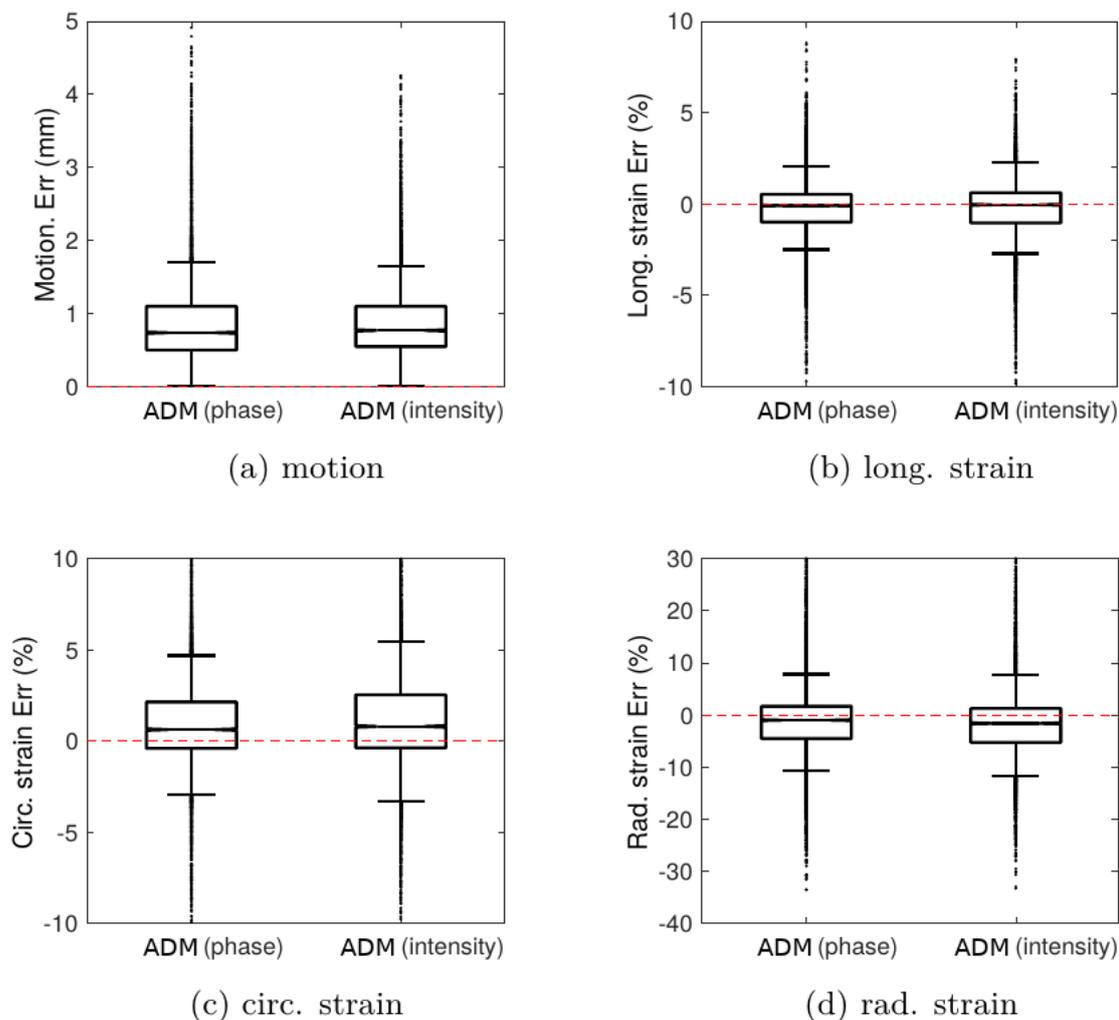


FIGURE 23 – Local motion and strain quantification errors obtained by the the proposed ADM method using either the image intensity or the phase information as the data attachment term involved in Eq. (29). Both methods were run on the 18 synthetic tagged MR datasets simulated in Sec. 4. For each cardiac phase, local motion and strains errors were computed for each 17 AHA segments and then averaged from all time instants, all segments and all virtual patients so to produce the given boxplots.

MR and US imaging. The radial regularization strategy developed in the ADM approach appears to be well suited since it allows a better radial strain estimation in most of the tested imaging modalities (both tagged MR and US). We also observed from Fig. 22 that a given motion tracking algorithm can perform differently across different modalities. For instance, the ADM method outperformed the sparse demons algorithm for radial strain estimation on both tagged MR and US but get clearly worst results on cine MR. This might be explained by the absence of local "markers" in cine MR imaging inside the myocardium, making the tracking in the middle of the myocardium highly influenced by regularization strategies. Contrary to AMD, sparse demons algorithm exploits a wider spatial regularization term beyond the myocardium region. Therefore, any information outside the myocardium can contribute to the tracking, making the sparse demons algorithm better suited for cine MR imaging.

Tagged MR is usually defined as the gold standard for cardiac motion and strain analysis. From the boxplots given in Fig. 22, it is interesting to note that this claim is also verified. Indeed, the results obtained by both the ADM and sparse demons algorithms are globally better when using tagged MR sequences. However, it is also interesting to see that the results obtained from the US sequences are not so far compared to the ones obtained from the tagged MR. This emphasizes the usefulness of US modality when the quality of the acquisition is high enough (which is the case for the template recordings that we used to generate the synthetic US sequences). In the futur, we plan to simulate and add typical artifacts (such as shadows or signal dropouts) in our US synthetic sequences to study their relative influence on the quality of motion and strain quantification compared to the corresponding performances on tagged MR from the same virtual patient.

Although the results presented in Fig. 22 constitutes a strong basis for motion and strain quantification analysis, additional experiments are still needed to make more relevant conclusions. For instance, it would have been interesting to evaluate the accuracy of the global indexes that can be extracted from each simulated modality, such as the left ventricle volumes, the ejection fraction and the myocardial mass. It would also have been important to study the performance of the tested methods according to the pathology. Unfortunately, due to the limited time and ressources, we were not able to conduct yet a more advanced study that could have harnessed the full potential of our synthetic dataset.

5.6 Conclusions

In this section, we developed a novel framework based on an anatomical deformable model to efficiently quantify myocardial motion field and strain at once. In the proposed formulation, myocardial motion tracking is regarded as a succession of image-to-image registration problem. To take advantage of the temporal coherence of the sequence, we adopted a frame-to-frame strategy. This framework was assessed through the synthetic US, tagged MR and cine MR sequences whose pipeline is described in Sec. 4. Thanks to this virtual dataset, the following insights were underlined :

- ADM and sparce demons algorithms provide better motion estimation and strain quantification on tagged MR and US compared to cine MR ;
- although the best results are obtained on tagged MR, ultrasound data allow close motion estimation and strain quantification when the quality of the acquisition is high enough ;
- a given motion tracking algorithm can perform differently across different modalities ;

5. *2nd contribution : anatomical deformable model for strain estimation in cardiac imaging*

- ADM appears to be the most interesting method for quantifying tagged MR, while the sparse demons algorithm seems to be better suited for cine MR and US imaging ;
- the radial regularization strategy developed in the ADM allows for a better radial strain estimation on both tagged MR and US imaging.

6 Summary - Conclusions

From Sec. 4 to 5 we presented our contributions on motion estimation and strain quantification in US, cine and tagged MR imaging. From the observation that a well-established validation of strain imaging was still needed in our community, we developed a novel pipeline specifically designed to simulate realistic synthetic US and MR image sequences from the same virtual patient. The proposed pipeline was used to create a dataset of 18 virtual patients including healthy and various pathological cases, *i.e.* ischemia and dyssynchrony. The aim of this benchmarked dataset was to provide the community with a fair and reproducible comparison of state-of-the-art methods for motion estimation and strain quantification. The following innovations were introduced in this study to generate cardiac sequences with high degree of realism :

- the combination of an E/M model with US/MR physical simulators to introduce the possibility of interacting with the image formation process ;
- the use of multimodal template sequences from the same patient to extract the most relevant information from each modality ;
- the introduction of a novel combinative warping strategy, with the goal of reducing motion artifacts that may occur in myocardium regions ;
- the contribution of the first unified framework which allows the generation of multimodal (US, cine and tagged MR) realistic synthetic sequences for the same virtual patient ;
- the use of a novel strategy based on a Gaussian weighting function to generate a smooth interface between the myocardium and the surrounding structures.

Based on the literature review we made in Sec. 2, we also developed a novel method, named anatomical deformable model (ADM), integrating the advantages demonstrated by several available strategies to efficiently quantify myocardial motion field and strain at once. The proposed ADM method is based on the following key aspects :

- the expression of cardiac deformations from the myocardial space. By doing so, the integration of any prior such as local motion or regularization will be efficiently adapted to the anatomy of the heart ;
- the use of a windows-weighted regression method to regularize the estimated motion and strain with low computational complexity ;
- the reinforcement of myocardial incompressibility by correcting strain values in the radial direction only.

Thanks to the proposed virtual dataset, ADM could be quantitatively compared to one of the most recent state-of-the-art methods, *i.e.* the sparse demons. The following conclusions/insights could then be drawn :

- ADM and sparse demons algorithms provide better motion estimation and strain quantification on tagged MR and US compared to cine MR ;
- although the best results are obtained on tagged MR, ultrasound data allow close motion estimation and strain quantification when the quality of the acquisition is high enough ;
- a given motion tracking algorithm can perform differently across different modalities ;
- ADM appears to be the most interesting method for quantifying tagged MR, while the

sparse demons algorithm seems to be better suited for cine MR and US imaging ;

- the radial regularization strategy developed in the ADM allows for a better radial strain estimation on both tagged MR and US imaging.

Based on the current state of the results we have obtained so far, it seems that there is still a need to improve motion estimation from cardiac images, although integrative models allow effective regularization of motion to better estimate myocardial strain values. From the very encouraging results we recently got in MRI and ultrasound cardiac segmentation, it appears natural to study the effectiveness of applying deep learning solutions for such a task and see if statistical learning approaches can be good candidates to also solve the problem of motion estimation in cardiac imaging.

My research on cardiac motion estimation was carried out through 1 PhD funded by an industrial grant (French CIFRE PhD with Philips Medisys compagny) and 1 postdoc funded by the french government (financed through the US-tagging ANR project of professor Hervé Liebgott). Results have led to 9 articles in international journals and 11 articles in international conferences.

Ultrasound image reconstruction : toward new imaging techniques

1 Introduction

As described in Sec. 1, echocardiography is one of the most common modality used in clinics. Among the reasons of its attractiveness, one can mention the fact that ultrasound is safe since no radiation is involved, real-time, portable and cheaper compared to other imaging devices such as MR and CT. In echocardiography, the real-time nature of ultrasound devices allows the investigation of both anatomical and functional information, making it possible to establish relevant diagnosis at patient bedside. In the last decade, many studies have been proposed in the literature to measure new clinical indices with high potential for the diagnosis of diseases thanks to the imaging of dynamic phenomena at high frame rates [52, 83, 275]. The key concept of these methods resides in the capacity of transmitting hundreds, or even thousands frames per second by rethinking the conventional ultrasound imaging process. Unfortunately, the increase of the frame rate deteriorates the quality of the reconstructed information and a trade-off between frame rate and information quality must be found.

There exists several ways to acquire ultrasound images at high frame rate. One way is to use the so-called parallel receive beamforming technique to generate several imaging beams in parallel thanks to the transmission of broader beams [239]. Such transmission can be achieved by reducing the transmit aperture [103, 276] and by using unfocused or defocused beams [101]. In this way, the frame rate can be increased with a factor equals to the number of parallel received beams. However, the increase of the number of parallel received beams must be followed by an increase of the width of the transmitted beams, resulting in a lower image quality (*i.e.* signal-to-noise ratio, contrast and resolution). In order to maintain high resolution while using fewer broad beams to illuminate the desired field of view, Synthetic Transmit Aperture (STA) imaging has been proposed [119]. A single element is used to transmit a spherical wave covering the full image view and all of the elements are used in reception to form a low resolution image. By coherently combining the images from successive transmissions, one may produce transmit dynamic focusing along each line of the final image.

Another way to increase the frame rate is to use multiple transmit beams, called Multi-Line Transmission (MLT). With this technique, several focused ultrasound pulses steered in different lateral directions are transmitted and the same number of image lines are beamformed simultaneously [157]. Using MLT, the frame rate can be M times faster, where M is the number of parallel pulses in each transmission. However, the interferences, also known as cross-talk between the simultaneous beams may create artifacts in the resulting image, for instance bright targets can appear in several locations in the image. Various methods have been proposed to reduce

the cross-talks by separating the beams either spatially or spectrally [59–61, 73], using beam transformation techniques [68], second harmonic signal [207] or various apodization schemes [260, 261].

In order to significantly reduce the number of transmissions, the concept of ultrafast ultrasound imaging has been proposed. This type of approach uses a smaller number of defocused ultrasound waves in transmission to insonify the entire medium [254]. Coherent compounding of the ultrasound image acquired from each transmission allows for a synthetic focusing in the full image, as in Synthetic Transmit Aperture imaging [119]. Several approaches were proposed and validated based on this concept with plane waves (PW) [48, 90, 154] and diverging waves (DW) insonifications [101, 191] for different applications, such as transient elastography [167] and cardiac imaging [101, 191]. Provost *et al.* demonstrated the feasibility of compounding DW using a sparse virtual array located behind the probe to achieve high frame rates for 3D cardiac imaging [210]. All these methods are based on the same scheme : *i*) use of PW or DW to insonify the whole medium for each firing; *ii*) the backscattered echoes are then measured and processed to reconstruct simultaneously all lines of the image of interest either in the Fourier domain [48, 90, 154] or in the space domain [101, 167, 191]. It has been shown in [90] that the computational complexity of the reconstruction process with PW insonifications performed in the space domain is higher than in the Fourier domain. As far as cardiac imaging is concerned, all the existing ultrafast methods rely on a sectorial configuration associated to DW, due to the limited anatomical access to the heart. Moreover, it has to be noted that all the DW-based proposed methods [101, 191, 210] perform reconstruction in the spatial domain, resulting in a higher computational complexity.

The remainder of the chapter is organized as follows.

- Sec. 2 : Review of state-of-the-art methods in ultrafast image reconstruction ;
- Sec. 3 : Presentation of a complete study to revisit the influence of the transmission scheme on plane-wave imaging system ;
- Sec. 4 : Overview of a novel formalism, named Ultrasound Fourier Slice Beamforming (USTB), that we developed to reconstruct ultrasound images from steered plane-waves ;
- Sec. 5 : Detailed description of a novel method dedicated to an efficient reconstruction of ultrasound sectorial images using diverging waves ;
- Sec. 6 : Conclusions of this chapter.

2 State-of-the-art in ultrafast image reconstruction

One of the advantage of PW/DW imaging techniques resides in their ability to reach high frame rate with controlled image quality by playing with the number of firings involved for the reconstruction of one image [203, 209]. This aspect motivated us to work with such a technique. A synthetic review of the classical ultrafast imaging techniques is proposed below.

Plane wave imaging

Lu *et al.* first proposed in the late 90s a Fourier-based method for high frame rate imaging with limited diffraction beams (PW being a particular case) [154, 155]. From the use of limited diffraction beams with normal incidence both in transmission and reception, the authors demonstrated that it is possible to reconstruct the Fourier spectrum of the object function from the

following equation :

$$k_z = k + \sqrt{k^2 - k_x^2}, \quad (1)$$

where (k_x, k_z) corresponds to the wavenumbers of the image to reconstruct and k is related to the temporal frequency of the received signal ($k = 2\pi f/c$, with c being the speed of sound).

Garcia *et al.* proposed a f-k migration method for 2D PW imaging based on the exploding reflector model (ERM) [90]. Their model assumes that all the reflectors in the medium explode simultaneously and become upward-emitting acoustic sources [92]. By using the virtual exploding sources, the ERM can reflect the actual two-way propagation of the acoustic wave accurately. By applying the Stolt's method to find these virtual sources and applying a spatial transformation, the actual scatterers' positions can thus be recovered. Based on this formalism, ultrafast imaging may be reconstructed from the transmission of a single PW with normal incidence thanks to the following relation :

$$k_z = 2 \sqrt{k^2 - k_x^2/2} \quad (2)$$

More recently, Montaldo *et al.* introduced a 2D time domain PW imaging method based on the principle of delay-and-sum (DAS) beamforming [167]. Contrary to the methods proposed by Lu and Garcia, this technique performs the reconstruction of PW imaging directly from the received raw data without the need to pass through the Fourier spectrum. The whole process of this method for a single PW is presented in Fig. 1. More particularly, a PW is transmitted into the medium by exciting the transducer elements simultaneously (Fig. 1a) and the backscattered echoes are acquired by each element at position x (Fig. 1b). The corresponding signal is named as $RF(x, t)$ in the sequel. By applying time delays on the raw RF signals and adding the resulting signals coherently, the whole image can be reconstructed (Fig. 1c). In the particular case of a steered PW with angle θ , the travel time to the point (x_s, z_s) and back to a transducer element placed in x can be expressed as [167] :

$$\tau(x, x_s, z_s) = \frac{1}{c}(\sin(\theta)x_s + \cos(\theta)z_s + \sqrt{(x_s - x)^2 + z_s^2}) \quad (3)$$

Each pixel of the image (x_s, z_s) is obtained by delaying the $RF(x, t)$ signals with $\tau(x, x_s, z_s)$ and adding coherently in the transducer direction x (see Eq.(2) in [167]) :

$$s(x_s, z_s) = \int_{x_s-a}^{x_s+a} RF(x, \tau(x, x_s, z_s))dx \quad (4)$$

The aperture $2a$ represents the elements that contribute to the signal and can be expressed by the F-number (see Eq.(3) in [167]) :

$$F = \frac{z}{2a} \quad (5)$$

where z represents the depth of the image. Ideally, the F-number is constant in the entire image.

Diverging wave imaging

More recently, Papadacci *et al.* proposed to adapt Montaldo's method to ultrafast imaging with DW in transmission for imaging the heart in 2D [191]. In this formalism, each transmitted DW is defined by the position of a virtual source located behind the probe and used to insonify the entire field of view. Hasegawa *et al.* also proposed a similar approach to perform high frame rate

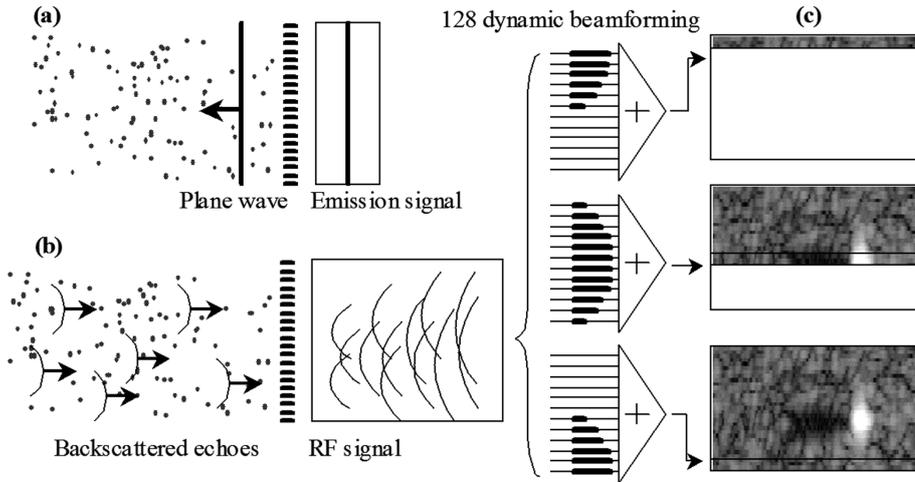


FIGURE 1 – (Figure 2 in [167]) Schematic representation of the single transmit PW method. (a) The ultrasonic array insonifies the medium using a PW transmission. (b) The backscattered RF signals are recorded by the transducer array. (c) The beamforming procedure consists in applying time delays laws and summations to the raw RF signals to focus in the receive mode. Contrary to standard ultrasonography, each line of the image is calculated using the same RF data set but with a different set of time delays.

imaging in 2D echocardiography, but using several diverging beams with different directions to cover the entire medium (15 transmits in [101]). The beamforming process of these methods is also based on the DAS principle. In particular, if a 2D DW is sent from a virtual source positioned at (x_v, z_v) , the time to travel to a point (x_s, z_s) in the medium and come back to a transducer placed in x can be written as :

$$\tau(x, x_s, z_s, x_v, z_v) = \frac{1}{c} (\sqrt{(x_s - x_v)^2 + (z_s - z_v)^2} + \sqrt{(x_s - x)^2 + z_s^2}) \quad (6)$$

Thanks to Eq. (6), a 2D sectorial image can be obtained with the same principle as the one used in PW imaging.

2.1 Principle of coherent compounding

The insonification with a single wave (PW or DW) provides the highest frame rate, but comes up with an image quality (defined in terms of resolution and contrast) intrinsically lower than the one obtained with classical multi-line focused beamforming. To overcome this limitation, spatial coherent compounding has been used in all the studies mentioned above to improve image quality. By using several PW with different angles (or several DW with different virtual source positions), synthetic focus is achieved in the whole image, as it is done in synthetic transmit aperture (STA) [119]. Fig. 2 shows the principle of coherent PW compounding. Steered PW are sent through the transducer and insonify the whole region of interest. An ultrasound image is computed from each single insonification. The individual image obtained from each PW has low quality. The coherent summation of these images creates synthetic focus throughout the image and allows to recover a high quality image. It has been shown in [167] that the synthetic focusing reached by coherent compounding is the same as in the standard focusing method for an adequate number of PW, allowing the derivation of images with the same level of quality.

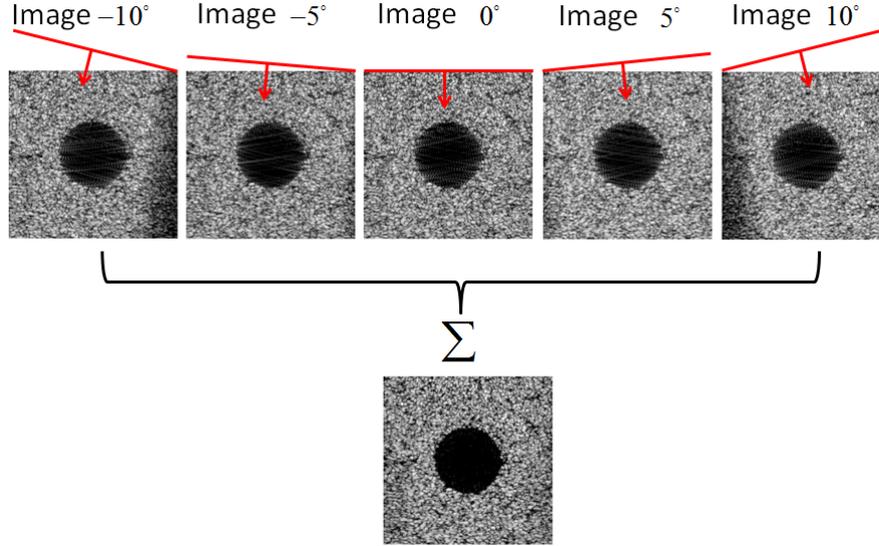


FIGURE 2 – Principle of coherent PW compounding. The medium is insonified with steered PW, and the images are reconstructed individually for each PW. The individual images obtained from each PW are low quality images. The coherent summation of these images creates synthetic focusing throughout the complete image and allows to recover a high quality image.

2.2 Motion compensation methods for ultrafast imaging

As described above, the final image produced by ultrafast imaging techniques (based on PW and DW insonifications) are dependent on the coherent summation of images obtained from several transmissions. However, when the scatterers move rapidly between two transmits, a time shift appears in the two successive received beamformed signals, resulting in a degradation of the image quality during coherent compounding [281]. When the motion is less than $\lambda/4$ (λ representing the wavelength), this time shift can be approximated by a phase shift in the received quadrature-demodulated (IQ) signals [62]. For moving tissues, this phase shift must be taken into account during the compounding process to decrease the image artifacts caused by motion. Several approaches have thus been proposed to tackle this problem in synthetic transmit aperture imaging (STAI) [87, 127, 180, 263, 290]. Recently, motion compensation methods adapted to coherent compounding with PW and DW insonifications were investigated in [62] and [203], respectively. In these works only radial motion (*i.e.* motion perpendicular to the transducer surface) is considered since lateral motion (*i.e.* motion parallel to the transducer surface) is expected to have a smaller influence on the image [281]. Both of these techniques are briefly described below.

Cross-correlation based motion compensation method for plane wave imaging

In order to reduce the motion artifacts on PW compounded images, Denarie *et al.* proposed a transmit sequence with angles alternatively negative and positive (*i.e.* transmit angle sequences $[-\alpha_N, \alpha_N, -\alpha_{N-1}, \alpha_{N-1}, \dots, -\alpha_1, \alpha_1, 0]$), which allows reducing the influence of motion while keeping the same contrast as the classical linear sequences [62]. However, the proposed transmit sequence cannot fully cancel the motion artifacts, so the same authors proposed to adapt the cross-correlation method to estimate the motion between each PW acquisition and correct it

before coherent compounding.

Doppler-based motion compensation for diverging wave imaging

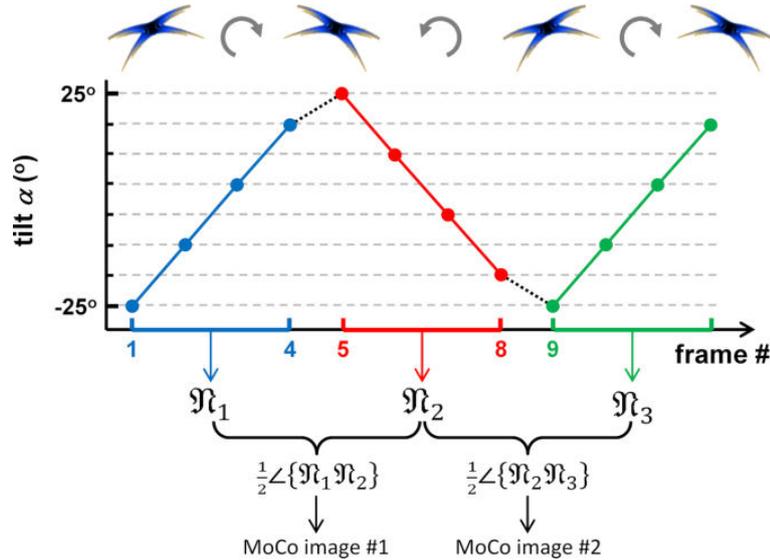


FIGURE 3 – (Fig. 3 in [203]) Triangle transmit sequences & motion estimation. To reduce the side effects due to sidelobes, a triangle sequence composed of 32 successive transmits is used. This figure shows an 8-transmit sequence for clarity. Slow-time autocorrelations were calculated in the ascending and descending stages, and their product was used to estimate the phase delays due to motion. The insets on the top represent the PSFs (point spread functions) with the apparent sidelobes. They rotate in clockwise then counter-clockwise directions.

In a recent paper, Poree *et al.* proposed a motion compensation (MoCo) strategy for high frame rate echocardiography using a sequence of DW in transmission [203]. To eliminate the artifacts due to the sidelobes, the authors proposed to use a triangle sequence as shown in Fig. 3. Two lag-one autocorrelations corresponding to the ascending and descending slopes of the triangle sequence were calculated and their product was used to estimate the phase delays due to motion. Finally, by compensating the displacements and phase rotations on each lower-resolution image and doing coherent summation, the final compounded image can be obtained.

3 Revisiting the influence of transmission scheme involved in ultrafast imaging with plane waves

In this section, a complete study is given to revisit the influence of the transmission scheme on the PW imaging system. Based on an analytical study of the problem, we consider a so-called efficient dichotomous transmission scheme for linear acquisition which allows improving the quality of the reconstructed images at a constant frame rate. The efficiency of the transmission dichotomous scheme is assessed through the use of the current state-of-the-art reconstruction algorithms for high frame rate imaging. The quality of the reconstructed images is evaluated in terms of resolution and contrast from both simulations and *in vitro* experiments. Comparison with the regular transmission scheme demonstrates the potential of the dichotomous approach which allows to divide by a factor of four the number of steered PW while maintaining the same image quality in the particular case of the reconstruction of static images.

3.1 Motivations

As discussed in the previous section, PW imaging in ultrasound is an intense area of research because of its capacity of reaching frame rate higher than a thousand of frames per second, leading to applications with strong potential in medical ultrasound [29, 77, 78, 93, 142, 167, 171, 208, 210, 221, 233, 268]. Contrary to the historical line-by-line focused beams approach, this technique is based on the reconstruction of a conventional B-mode image from the transmission of one or several steered PW. Although PW imaging appears to be attractive, the images obtained using a single PW suffer from an intrinsic lower quality (defined in terms of resolution and contrast) as compared to the conventional multi-lines focused beamforming technique, especially for large depths. Coherent compounding based on the transmission of N steered PW with different angles corresponds to the best-established approach to overcome this limitation [167]. The interest of such technique resides in its ability to reproduce a synthetic focus on the full image in transmission, as it is done in synthetic transmit aperture [119]. The quality of the coherent summation was first assessed from static targets on numerical simulation and/or *in vitro* physical phantoms [48, 90, 167]. The conventional scheme proposed in the corresponding papers involved a set of steered PW transmitted with an increment angle of 1° . For instance, conventional PW imaging techniques based on the transmission of 5 steered PW classically involves angles of values $(-2^\circ, -1^\circ, 0^\circ, 1^\circ, 2^\circ)$. From this framework, it has been observed in most of the studies that the image quality monotonically increases with the number of transmitted steered PW up to a convergence state [90, 167]. One of the main interest of PW imaging is its capacity in imaging dynamic phenomena at high frame rates. In the case of moving targets, the perfect coherent summation is no longer achieved [281]. The influence of the maximum transmitted angle α_{max} and the number of steered PW on the image quality were analytically investigated in [62]. An alternated transmit scheme with motion compensation was then proposed to reduce the influence of motion on PW compounded image. In this section, we propose to revisit the influence of the transmission scheme in the case of PW imaging. Since most of the new concepts introduced in this manuscript are validated from static phantoms (both in simulation and from *in vitro* experiments), we focused our attention on the optimization of the transmission scheme in the particular case of non-moving object.

3.2 Methodology

3.2.1 Principles

Let's start from the explicit expression of a monochromatic PW $p_c^i(x, z, t)$ with a steering angle α_i [167] :

$$p_c^i(x, z, t) = p_0 \exp(j(xk_x^i + zk_z^i - \omega t)) \quad (7)$$

where $k_x^i = k_0 \sin \alpha_i$, $k_z^i = k_0 \cos \alpha_i$, $k_0 = 2\pi/\lambda$ (λ corresponding to the conventional wavelength) and $\omega = 2\pi f_0$. In the particular case of a set of N PW steered in a symmetric manner with a constant increment angle $d\alpha$, α_i can be reformulated as $\alpha_i = i d\alpha$ with $i \in [-(N-1)/2, (N-1)/2]$, leading to the following relation for small values of α_i :

$$k_x^i = \frac{2\pi}{\lambda} \sin(i d\alpha) \approx \frac{2\pi}{\lambda} i d\alpha. \quad (8)$$

As demonstrated in [167], the image of a point (x_f, z_f) from compounding scheme is the same as the one obtained by illuminating the medium with a unique wave :

$$p_c(x', z', t) = \sum_{i=-(N-1)/2}^{(N-1)/2} p_0 \exp(j(x'k_x^i + z'k_z^i - \omega t)), \quad (9)$$

where $x' = x - x_f$ and $z' = z - z_f$ are the coordinates centered at the focal point (x_f, z_f) . At the focal depth $z = z_f$, the evolution of the synthetic pressure field along the x-axis is then given by :

$$p_c(x', t) = \sum_{i=-(N-1)/2}^{(N-1)/2} p_0 \exp(j(x'k_x^i - \omega t)). \quad (10)$$

By using the relation given in Eq. (8) and the standard geometrical series relations, the above equation can be simplified as :

$$|p_c(x', t)| = |p_0| \left| \frac{\sin(k_0 d\alpha x' N/2)}{\sin(k_0 d\alpha x'/2)} \right| \quad (11)$$

The above expression allows investigating the dimension of the main lobe and the positions of the side lobes in the monochromatic case at the center frequency and at the focal depth. Indeed, Eq. (11) vanishes for

$$x' = \frac{\lambda}{Nd\alpha} l, \quad l \in Z^* \quad (12)$$

The main lobe width is proportional to the distance between the first zeros of the transverse pressure function (*i.e.* the distance between $l = -1$ and 1 in Eq. (12)), thus the width W of the main lobe corresponding to a coherent summation of N steered PW ($N > 1$) is in the order of :

$$W = 2 \frac{\lambda}{Nd\alpha} = \frac{N-1}{N} \frac{\lambda}{\alpha_{max}} \approx \frac{\lambda}{\alpha_{max}} \quad (13)$$

where $\alpha_{max} = \frac{N-1}{2} d\alpha$ represents the maximum transmitted angle. Eq. (13) shows that the width of the main lobe, which is directly linked to the lateral resolution of the system, is uniquely fixed by two parameters : the wavelength of the probe and the maximum transmitted angle α_{max} .

Thus, by maintaining the maximum angle α_{max} for varying number of steered PW involved in transmission, the corresponding reconstructed images will share the same resolution.

3.2.2 The dichotomous transmission scheme

For conventional imaging, the focused beam profile can be expressed as [62] :

$$|p(x', z_f, t)| = |p_0| \left| \frac{\sin(k_0 x' L / (2z_f))}{\sin(k_0 x' \Delta x / (2z_f))} \right| \quad (14)$$

where L is the size of the aperture, Δx is the pitch of the linear array. To achieve the equality of Eq. (11) and (14), the number of steered PW N and the angle increment $d\alpha$ have to satisfy the following relation :

$$Nd\alpha = \frac{L}{z_f} \quad (15)$$

By replacing $d\alpha$ in the above expression using the relation given in Eq. (13), the maximum transmitted angle α_{max} can be reformulated as

$$\alpha_{max} = \frac{N-1}{N} \frac{L}{2z_f} \approx \frac{L}{2z_f} \quad (16)$$

This last expression shows that the maximum transmitted angle α_{max} can be determined by the aperture size L and a chosen focused depth z_f . Based on Eq. (13) and (16), there can be several options to choose the angles in between, but the most straightforward is to distribute them equally in between the extreme values. We named this approach the dichotomous scheme thereafter and the corresponding algorithm is described below.

1. We first fix the maximum depth z_f that we want to image.
2. Thanks to Eq. (16), we then compute the optimal maximum angle α_{max} that we use whatever the number of steered PW involved in compounding.
3. The values of the angles of the steered PW used in transmission are then equally distributed from $-\alpha_{max}$ to α_{max} .

3.3 Experiments

3.3.1 Simulation

Based on the previous analytical study, we proposed to quantify from simulation the influence of the increment angle $d\alpha$ together with the number of steered PW used in compounding on the resolution of the transmitted field. To this aim, simulations of the synthetic acoustic field (acoustic field from each transmit summed with spatial coherent compounding) were performed to assess the system quality in transmission. A standard linear-array probe of 128 elements with 0.19-mm pitch, 13-mm height, a center frequency of 5-MHz (100% bandwidth), and a 23-mm elevation focus was implemented using Field II [120]. The sampling frequency was set to 100 MHz. The acoustic field transmitted by each steered PW was also calculated using Field II. Because this part of the study only focuses on the transmission side, it is worth noting that its

conclusions should benefit to any reconstruction methods based on PW imaging. This point will be investigated in the next section.

Influence of the increment angle $d\alpha$

Fig. 4 shows the transmitted synthetic pressure field obtained in the particular case of 3 steered PW with values of steered angles equal to $(-\alpha_s, 0, \alpha_s)$, α_s varying from 1° to 21° . In the particular case of 3 steered PW, α_s corresponds both to the increment angle $d\alpha$ and the maximum transmitted angle α_{max} . $\alpha_s = 1^\circ$ corresponds to the conventional scheme used in [48, 90, 167]. Coherent summation was performed at 30-mm depth and for two different lateral positions of the field of view : at the center (left-hand side of Fig. 4) and close to the edge of the probe (right-hand side of Fig. 4). The first row of Fig. 4 displays the lateral resolution, measured as the width of the main lobe at -6 dB, at 30-mm depth and for increasing values of α_s . The second row of Fig. 4 shows the normalized synthetic pressure distribution along the lateral position at 30-mm for α_s equals to 1° , 5° , 10° and 15° , while the last row displays these pressure distributions in the entire space. From Fig. 4(a) and 4(b), it can be seen that the lateral resolution improves with the increase of the value of the steered angle, which is in accordance with Eq. (13). For a steered angle greater or equal to 10° , a lateral resolution smaller than 1-mm is obtained. However, large values of α_s increase the magnitude of the grating lobes near the probe edges (Fig. 4(d)) because of the large $d\alpha$, which results in a poor resolution after 12° (Fig. 4(b)) and thus produced undesirable border effects in terms of imaging system. Therefore, an optimal steered angle α_{opt} around 10° appears to be a good trade-off between a good resolution (lateral resolution around 0.8-mm) and a small magnitude of the grating lobe. This value, fixed thanks to the above analytical study from a given probe settings, was used in the rest of the experiments.

Influence of the number of steered PW

Fig. 5 displays the evolution of the resolution with respect to the number of steered PW used in transmission computed from simulation of the synthetic acoustic field. The lateral resolution was measured at 30-mm depth in the center of the image. In this experiment, the conventional and the dichotomous scheme were compared. For the regular scheme, the angle increment is fixed to 1° , implying thus that the maximum angle increases with the number of steered PW. For the second approach, we used the optimal value $\alpha_s = 10^\circ$ derived in the particular case of 3 steered PW and then increased the number of steered PW uniformly from $-\alpha_s$ to α_s . For instance, in the particular case of 5 steered PW, the values of the involved steered angles are thus equal to $(-\alpha_s, -\alpha_s/2, 0, \alpha_s/2, \alpha_s)$, *i.e.* $(-10^\circ, -5^\circ, 0^\circ, 5^\circ, 10^\circ)$. From this figure, it can be observed that the lateral resolution improves with the increase of the number of steered PW in the case of the conventional approach (from 8-mm to 1-mm), which is in accordance with Eq. (13). Indeed, for a constant value of $d\alpha$, the width of the main lobe decreases for increasing values of the number of steered PW, which results in an improvement of the lateral resolution. On the opposite, the dichotomous transmission scheme yields a quasi constant resolution around 1-mm whatever the number of the involved steered PW, which illustrates the strong interest of such approach. Moreover, the small decrease of the lateral resolution after the compounding of 3 steered PW can also be explained from Eq. (13). Indeed, it can be analytically observed that the width of the main lobe slightly increases with the number of steered PW up to a convergence value equals to λ / α_s , which is in accordance with the results observed in Fig. 5.

Fig. 6 gives the normalized synthetic pressures computed along the lateral position at 30-

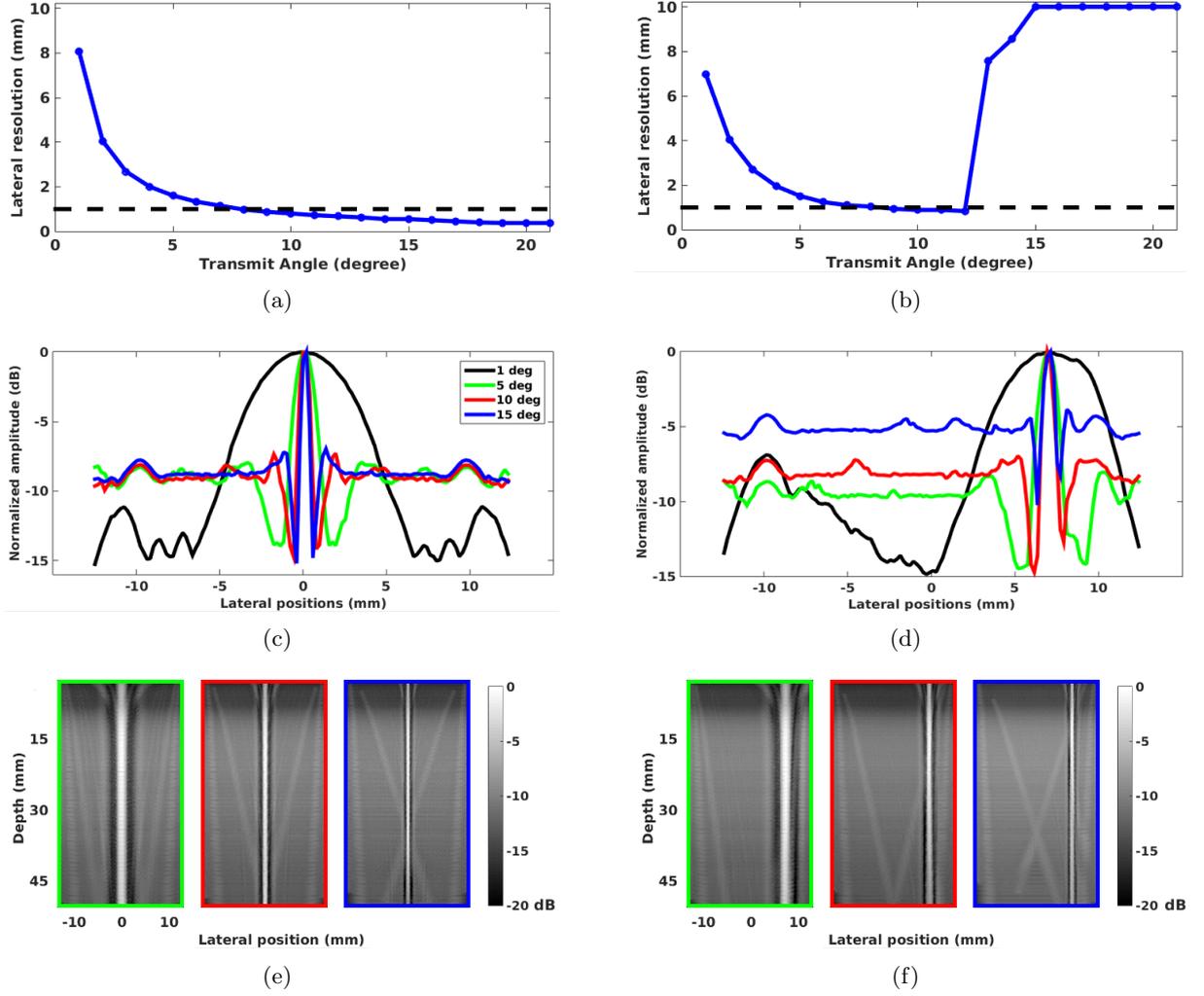


FIGURE 4 – Effect of the value of the steered angle α_s on the lateral resolution using synthetic acoustic field. Each pressure is compounded with 3 steered PW ($-\alpha_s, 0, \alpha_s$) at 30-mm depth on the central axis (left column) and on the edge of the probe (right column). (a)-(b) Lateral resolution as a function of α_s . The black dashed line represents the limit of lateral resolution at 1 mm. (c)-(d) Normalized synthetic pressure computed along lateral positions for different values of α_s ($1^\circ, 5^\circ, 10^\circ, 15^\circ$). (e)-(f) 2-D normalized synthetic pressure distribution obtained for the different configurations of α_s with $5^\circ, 10^\circ$ and 15° .

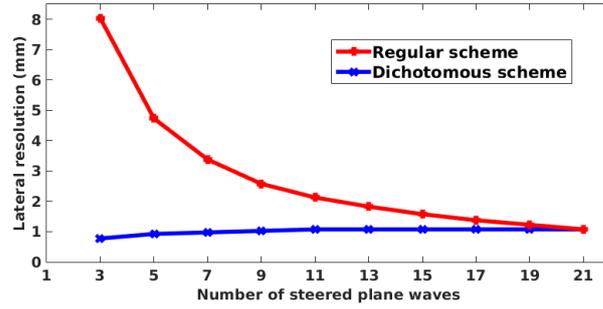


FIGURE 5 – Influence of the number of transmitted steered PW on the lateral resolution for the conventional and the dichotomous scheme.

mm depth using the two transmission schemes with increasing number of steered PW. For both methods, it can be seen that the grating lobes decrease with the number of steered PW, which should have an impact on the contrast. For the dichotomous scheme, this figure also reveals the stable behavior of the main lobe (lateral resolution), the increase of the transmitted steered PW having mostly an impact on the reduction of the grating lobes effects.

3.3.2 *In vitro* experiments

We evaluated the performance of the dichotomous and the conventional transmission schemes associated with the current state-of-the-art reconstructed methods for PW imaging, *i.e.* the methods of Lu [154], Garcia [90] and Montaldo [167]. The image quality was assessed experimentally using an ultrasound phantom (CIRS model : 054GS) [1]. Fig. 7 displays the schematic diagram of the corresponding phantom along with the two imaging planes involved in this experiment. A standard linear array probe (128 elements, center frequency of 7.8 MHz, transmitted frequency of 5 MHz, Verasonics - L12-5-50 mm) with the same characteristics as the one used in Sec. 3.3.1 was interfaced with a Verasonics system to image the phantom. The imaging depth was set to 50 mm. The number of transmitted steered PW varied from 1 to 21, providing a frame rate between 15400 and 733 fps, respectively. No apodization in transmission and reception was used for the entire acquisitions. The received raw-data were processed using the three state-of-the-art techniques for PW imaging described in Sec. 2. The two different transmission schemes presented in Sec. 3.3.1 were investigated. For the regular scheme, the angle increment is fixed to 1° while for the dichotomous scheme, the optimal steered angle α_s was set to 10° . The beamformed RF data were Hilbert transformed and normalized to get the envelope images. The corresponding images were then gamma-compressed using $\gamma = 0.3$ as in [90] and finally converted to 8-bit grayscale to get the B-mode images.

Lateral resolution

Lateral resolution was first investigated from the imaging plane n^o2 given in Fig. 7 (acquisition centered on the 0.1-mm nylon monofilament targets). The corresponding values were measured as the width at -6 dB of the Point Spread Function (PSF) corresponding to the two points present at 20 mm and 40 mm in the image. Fig. 8 investigates the improvement of the lateral resolution with the number of steered PW for the different reconstruction techniques and for different depths. First, it can be seen that the three methods produce similar image quality both for a single PW (Fig. 8a) or using 21 steered PW with coherent compounding (Fig. 8b). These observations are confirmed by the two graphs provided in Fig. 8c which display the evolution

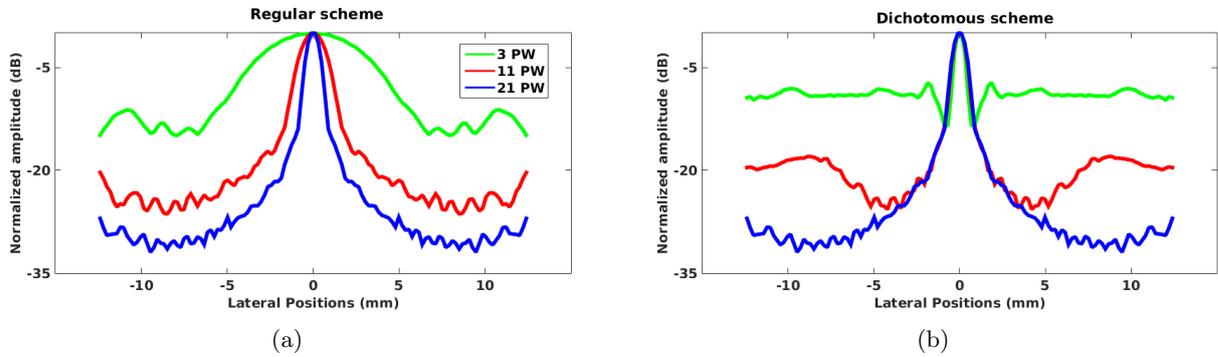


FIGURE 6 – Normalized synthetic pressure computed along the lateral position at 30-mm depth in the center of the image with varying numbers of steered PW (3, 11 and 21) using (a) the regular approach (*e.g.* 3 steered PW : $\{-1^\circ, 0^\circ, 1^\circ\}$) and (b) the proposed dichotomous scheme (*e.g.* 3 steered PW : $\{-10^\circ, 0^\circ, 10^\circ\}$).

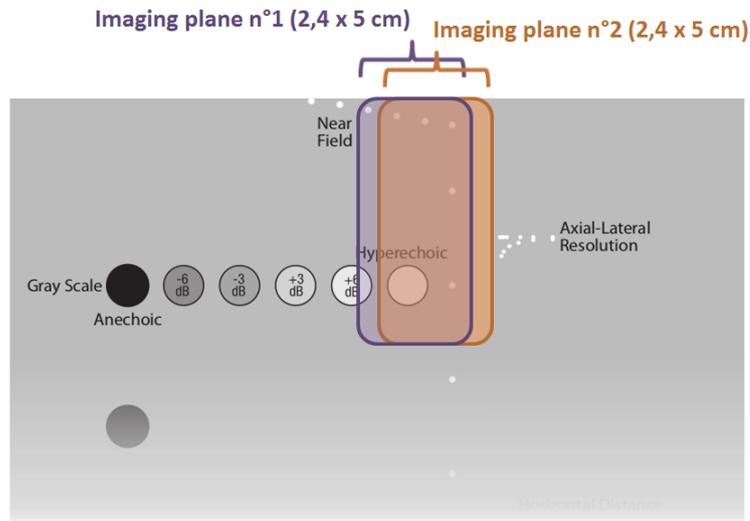


FIGURE 7 – Schematic diagram of the CIRS tissue-mimicking phantom (model : 054GS) with the corresponding imaging planes used in the experiments

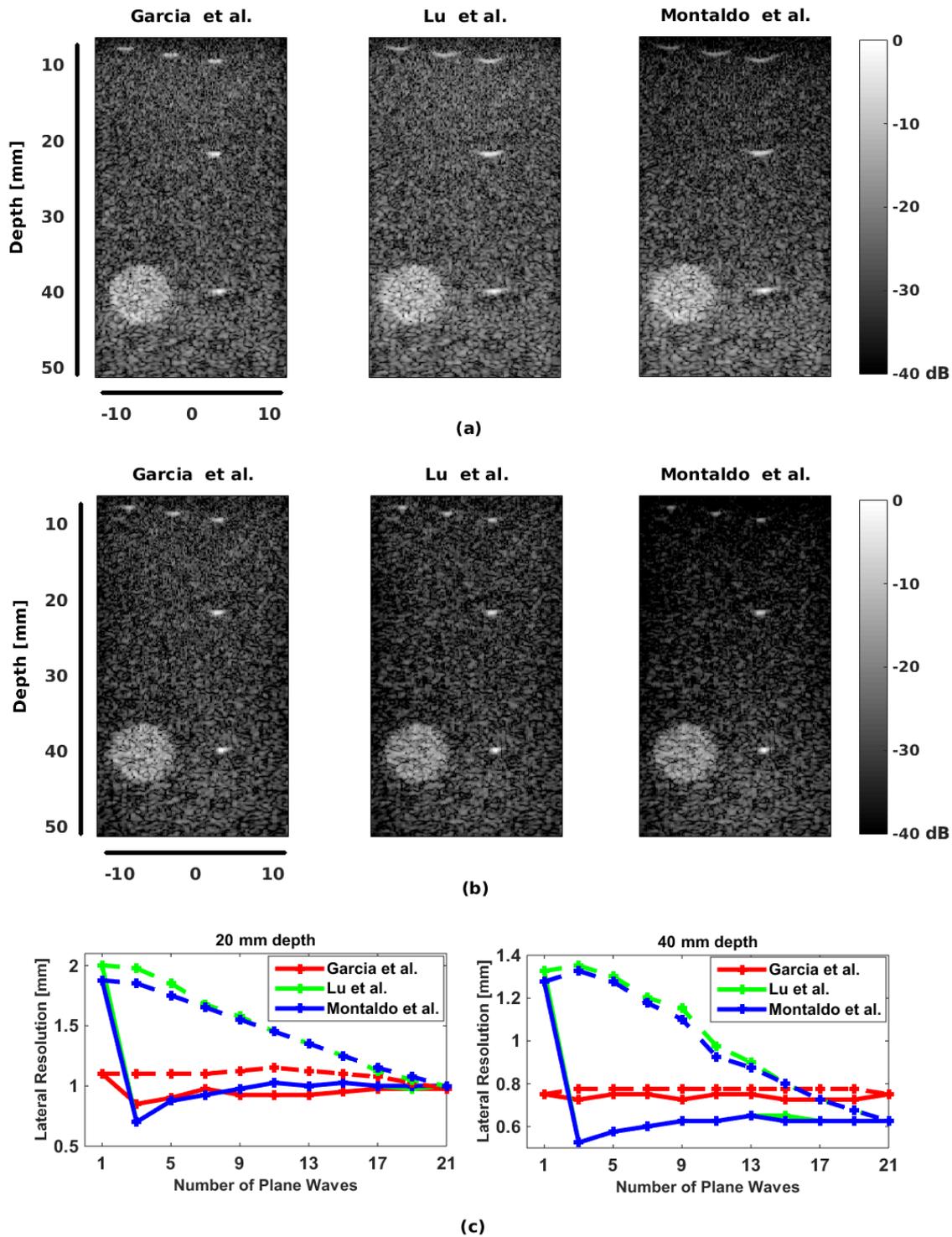


FIGURE 8 – B-mode images of a CIRS phantom (model : 054GS) reconstructed from the three reconstructed methods using (a) 1-PW and (b) 21 steered PW with coherent compounding. (c) Lateral resolution measurements as a function of the number of steered PW. The solid lines correspond the proposed dichotomous transmission scheme while the dashed lines correspond to the regular scheme.

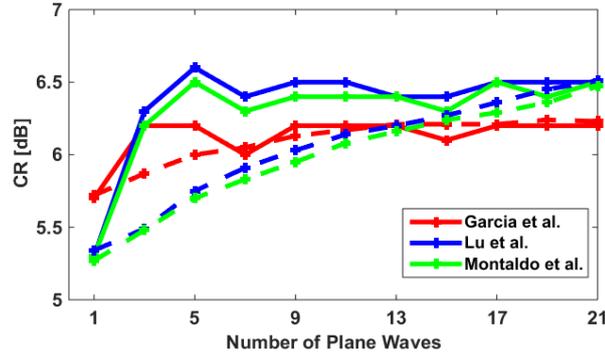


FIGURE 9 – Contrast ratio CR as a function of the number of steered PW for the three reconstructed methods. The solid lines correspond the proposed dichotomous transmission scheme while the dashed lines correspond to the regular scheme.

of the lateral resolutions measured at 20 and 40 mm with respect to the number of steered PW. Indeed, the lateral resolution derived from the dichotomous scheme improves quickly for 3 steered PW and then tends to stabilize to the optimal value whatever the reconstructed method, *i.e.* around 0.9-mm at 20-mm depth and between 0.6 and 0.8-mm at 40-mm depth, which is in coherence with what has been previously observed from simulations. The dichotomous scheme reaches the optimal lateral resolution values for the transmission of only 3 steered PW (frame rate of 5133 fps), while the conventional scheme (dashed lines) reaches the same values after the transmission of 21 steered PW (frame rate of 733 fps).

Image contrast

The image contrast was then investigated from the imaging plane n°1 given in Fig. 7 (acquisition centered on the hyperechoic target). Fig. 9 shows the improvement of the image contrast with respect to the number of steered PW for the three reconstruction methods. As for the lateral resolution case, the use of the dichotomous scheme in transmission (solid lines) allows a significant improvement of the image quality compared to the conventional approach (dashed lines) at a constant frame rate. Indeed, the dichotomous scheme reaches the optimal CR values for the transmission of only 5 steered PW (frame rate of 3080 fps), while the conventional approach reaches the same values after the transmission of 21 steered PW (frame rate of 733 fps). Moreover, it is also interesting to note that the reached optimal CR scores are close whatever the method, varying from 6 to 6.7-dB.

3.4 Conclusions

We revisited in this study the influence of the transmission scheme involved in PW imaging with coherent compounding. In particular, we investigated the importance of the choice of the steered angles on the quality of the reconstructed image in terms of resolution and contrast. Based on the analytical analysis of the synthetic pressure, we showed that the use of a dichotomous transmission scheme improves significantly the quality of the reconstructed B-mode images at constant frame rates in the particular case of static images. Experimental comparisons with the conventional transmission scheme reveal the potential of the dichotomous scheme in dividing by a factor of four the number of steered PW needed to reconstruct same quality of images, the optimal number of transmitted waves varying from 21 steered PW using the conventional approach to 5 steered PW with the dichotomous solution.

4 1st contribution : Ultrasound Fourier Slice Beamforming

In this section we present a novel formalism, named Ultrasound Fourier Slice Beamforming (UFSB), that we developed to reconstruct ultrasound images from steered plane-waves. The interest of such approach is to create strong links between ultrasound image formation with other modalities such as Computed Tomography (CT). Indeed, the principle of our method is to express the image beamforming through the Fourier domain where the samples are radially distributed. By doing so, standard methods such as the Fourier slice theorem can be used to easily reconstruct ultrasound images. This theory was validated through both simulation and in vivo acquisitions. This work has been the subject of several publications [35, 295].

4.1 Motivations

The principle of the existing ultrafast Fourier-based methods introduced in Sec. 2 can be summarized as follows. One PW is first emitted. Backscattered echoes are then measured (Fig. 10a). 2D-Fourier transform is applied on the received RF raw-data (Fig. 10b). Finally, a remapping function is used to project the Fourier transform of the received echoes to the k -space corresponding to the Fourier transform of the image to reconstruct (Fig. 10c). By taking the inverse 2D Fourier transform, the final ultrasound image is then reconstructed (Fig. 10d). One important shared property of the previous Fourier-based techniques is that the Fourier spectrum of the object is sampled along the k_z axis direction with a lateral step proportional to the inverse of the pitch (distance between two consecutive active elements). The Fourier slice theorem, which is based on the collection of projections of the image to reconstruct from many different directions, has been widely used in CT [108]. The theorem states that the one dimensional Fourier transform of a projection of an image provides a slice of the 2D Fourier transform of the image. Performing projections along different angles thus allows to radially sample the full spectrum of the image. By doing the inverse Fourier transform of the image spectrum, the image can be reconstructed. In CT systems, a beam of X-ray is created to pass through the tissue. A detector placed on the opposite side measures the corresponding projection along one direction. By rotating the emitter and detector for several angles, the full k -space of the image can be recovered. Unlike CT, ultrasound systems use the same transducer array to both transmit and receive the ultrasound wave. The use of the Fourier slice imaging technique in ultrasound is thus not straightforward and an adaptation of the projection reconstruction scheme to the specificity of ultrasound image formation is thus required.

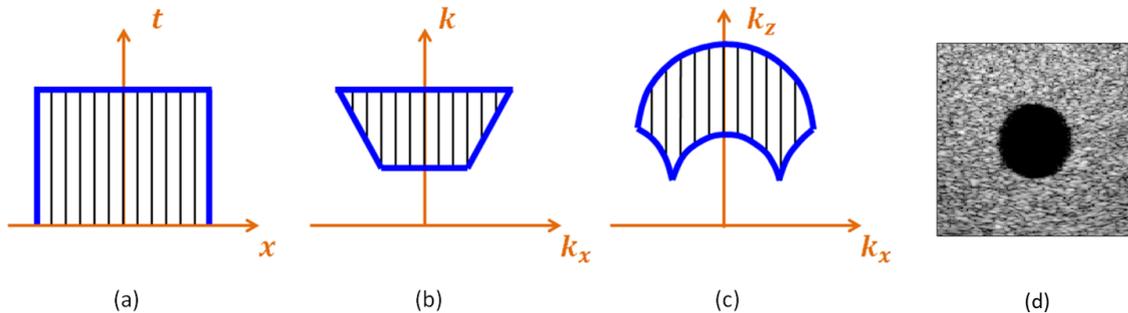


FIGURE 10 – Reconstruction procedure for conventional Fourier-based method. (a) the received raw-data in (x, t) space; (b) the 2D Fourier transform of the raw-data; (c) the remapped object spectrum; (d) the reconstructed image.

4.2 Methodology

4.2.1 UFSB : single plane wave imaging

Steered plane wave modeling

Ultrasound Fourier slice imaging is based on the use of steered PW both in emission and in reception. Let $\phi_{\mathbf{n}}(\mathbf{x}, t)$ be a plane-wave steered in the direction defined by \mathbf{n} . The scalar pressure field of this plane-wave can be defined as :

$$\phi_{\mathbf{n}}(\mathbf{x}, t) = \phi(ct - \mathbf{n} \cdot \mathbf{x}) \quad (17)$$

where $\mathbf{x} = (x, z)$ is a point in the medium with x is the lateral axis and z is the axial axis and ϕ represents the wave shape, which can thus be rewritten as :

$$\phi_{\mathbf{n}}(\mathbf{x}, t) = \phi(ct) * \delta(ct - \mathbf{n} \cdot \mathbf{x}) = \phi * \delta(ct - \mathbf{n} \cdot \mathbf{x}) \quad (18)$$

The wave front of a propagating field corresponds to the location, at a particular time instant t_0 , where the field value is the same, *i.e.* \mathbf{x} verifying $\phi_{\mathbf{n}}(\mathbf{x}, t_0) = cst$. In the particular case of steered plane-wave, thanks to Eq. (18), the corresponding wave front can be easily defined as the location where $\delta(ct_0 - \mathbf{n} \cdot \mathbf{x}) = 0$, *i.e.* $\mathbf{n} \cdot \mathbf{x} = ct_0$ which corresponds to a line perpendicular to the direction given by \mathbf{n} and passing through the point $ct_0\mathbf{n}$, as illustrated in Fig. 11.

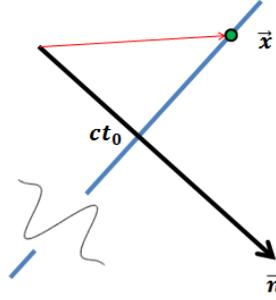


FIGURE 11 – Illustration of a wavefront of a steered PW with direction \mathbf{n} . All the \mathbf{x} in the blue line has the same wavefront.

Ultrasound image formation

We focus now on finding a relationship between the spectrum of the image to reconstruct and the received raw-data using a horizontal PW in transmission and steered PW in reception, as illustrated in Fig. 12. In particular, a linear array is excited to generate a single PW with direction \mathbf{n}_e to insonify the medium and is also used to receive the backscattered echoes coherently as a steered PW in the direction \mathbf{n}_r . The summation of the output signal of all the elements in such a system can be modeled as follows :

$$\begin{aligned} s(t) &= \sum_{Z^*} s(x_i, t) \\ &\approx \int_D m(\mathbf{x}) \cdot \phi_{\mathbf{n}_e}(\mathbf{x}) * \phi_{\mathbf{n}_r}(\mathbf{x}) dx \\ &\approx \int_D m(\mathbf{x}) \cdot \phi_{er} * \delta(ct - \mathbf{n}_{er} \cdot \mathbf{x}) dx \end{aligned} \quad (19)$$

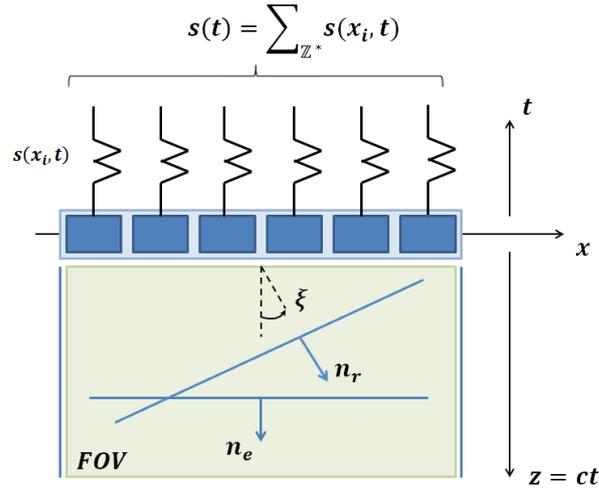


FIGURE 12 – Illustration of the image formation introduced in UFSB theory

where Z^* corresponds to the restriction of the element positions of the probe. $*$ represents the temporal convolution product and $\mathbf{n}_{er} = \mathbf{n}_e + \mathbf{n}_r$. ϕ_{er} corresponds to the convolution between the impulse responses of the emitted field ϕ_e and the received field ϕ_r . D corresponds to the transducer width and $m(\cdot)$ is a function that characterizes both the spatial distribution and the backscattered amplitude of the different scatterers present in the insonified medium. Typically, $m(\cdot)$ can be defined as :

$$m(\mathbf{x}) = \sum_{\{i,j\} \in FOV} a(i,j) \delta(x-i) \delta(z-j) \quad (20)$$

where $a(i, j)$ is the reflection coefficient of the scatterer positioned at (i, j) , FOV corresponds to the area insonified by the transmitted wave.

Ultrasound Fourier Slice Beamforming concept

The temporal Fourier transform $\mathcal{F}_t(\cdot)$ of signal $s(t)$ leads to the following relation :

$$F_t(s(t)) = \int_{-\infty}^{+\infty} \left[\int_D m(\mathbf{x}) \cdot \phi_{er} * \delta(ct - \mathbf{n}_{er} \cdot \mathbf{x}) dx \right] e^{-j2\pi ft} dt \quad (21)$$

Given the fundamental relations $z = ct$ and $k = \frac{2\pi f}{c}$, where z and k corresponds to the spatial distance and the wavenumber defined in the direction given by \mathbf{n}_{er} , respectively, Eq. (21) can be rewritten as :

$$\begin{aligned}
F_t(s(t)) &= \frac{1}{c} \int_{-\infty}^{+\infty} \left[\int_D m(\mathbf{x}) \cdot \phi_{er} * \delta(z - \mathbf{n}_{er} \cdot \mathbf{x}) dx \right] e^{-jkz} dz \\
&= \frac{1}{c} \int_D \left[\int_{-\infty}^{+\infty} m(\mathbf{x}) \cdot \phi_{er} * \delta(z - \mathbf{n}_{er} \cdot \mathbf{x}) e^{-jkz} dz \right] dx \\
&= \frac{1}{c} \int_D \left[\int_{-\infty}^{+\infty} m(\mathbf{x}) \cdot \phi_{er} e^{-jk(\mathbf{n}_{er} \cdot \mathbf{x})} dz \right] dx \\
&= \frac{1}{c} \iint_{FOV} m(\mathbf{x}) \cdot \phi_{er} e^{-jk(\mathbf{n}_{er} \cdot \mathbf{x})} dx dz
\end{aligned} \tag{22}$$

Since $m(\mathbf{x})$ is zero out of the field of view (FOV), we use FOV to represent the limit of integration in the last step of Eq. (22). In the special case where $\mathbf{n}_e = (0, 1)^T$ and $\mathbf{n}_r = (\sin(\xi_i), \cos(\xi_i))^T$, Eq. (22) can be rewritten as :

$$\begin{aligned}
F_t(s(t)) &= \frac{1}{c} \iint_{FOV} m(\mathbf{x}) \cdot \phi_{er} e^{-jk(\sin(\xi_i)x + (1 + \cos(\xi_i))z)} dx dz, \\
&= \frac{1}{c} \iint_{FOV} m(\mathbf{x}) \cdot \phi_{er} e^{-j(k_x x + k_z z)} dx dz, \\
&= \frac{1}{c} F_{2D} \left[m(\mathbf{x}) \cdot \phi_{er} \right] (k_x, k_z)
\end{aligned} \tag{23}$$

where F_{2D} is the 2D Fourier transform operator. According to the Fourier slice theorem, Eq. (23) can be interpreted as the 2D spatial Fourier transform of the image $(m \cdot \phi_{er})(\cdot)$ restricted to the line of direction \mathbf{n}_{er} , with the following fundamental relations :

$$\begin{cases} k_x = k \sin(\xi_i) \\ k_z = k (1 + \cos(\xi_i)) \end{cases} \tag{24}$$

and

$$\begin{cases} k = \frac{k_x^2 + k_z^2}{2k_z} \\ \xi_i = \arctan\left(\frac{k_x}{k_z - k}\right) = \arctan\left(\frac{2k_x k_z}{k_z^2 - k_x^2}\right) \end{cases} \tag{25}$$

As a consequence, in the particular case where the emitted field is a PW perpendicular to the probe and the received field is supposed to be a steered PW with angle ξ_i , Eq. (22)-(25) show that the temporal Fourier transform of the received signal is equal to a radial line of angle $\theta_i = f(\xi_i)$ (with $f(\cdot) = \arctan(\sin(\cdot)/(1 + \cos(\cdot)))$) in the corresponding k-space domain. By simply playing with different delay strategies applied on the received signals, we are thus able, for only one emitted PW, to radially and densely recover the Fourier space of the object and thus reconstruct an ultrasound image with high frame rate. The acquisition scheme corresponding to the proposed approach is summarized in Fig. 13.

4.3 UFSB : compounding imaging from steered plane waves

It has been shown in [90, 154, 167] that the image quality can be improved by coherently compounding the images obtained from several steered PW. To be able to perform coherent PW

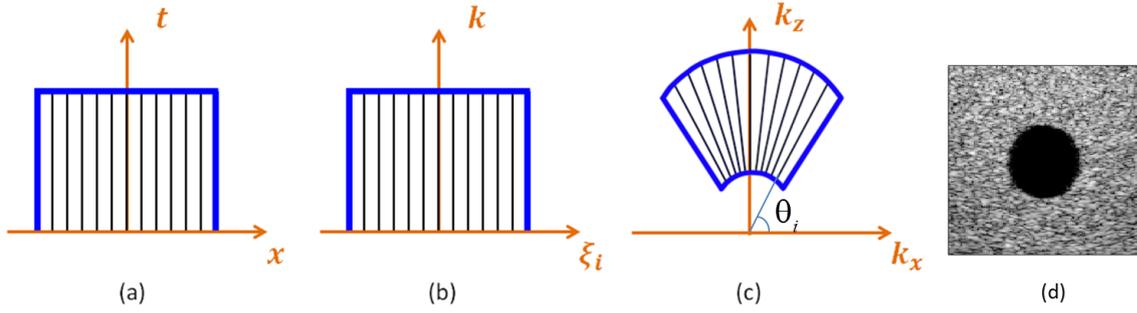


FIGURE 13 – Reconstruction procedure with the proposed method. (a) the received raw-data in (x, t) space; (b) the 1D Fourier transform of all the projections of the raw-data at different angles ξ_i ; (c) the remapped object spectrum; (d) the reconstructed image.

compounding with our method, it must be ensured that the UFSB technique can be adapted to a steered PW in transmission as well. For a steered PW with angle φ_j in transmission, which means the corresponding direction $\mathbf{n}_e = (\sin(\varphi_j), \cos(\varphi_j))^T$, the following relations can be easily derived :

$$\begin{cases} k_x = k (\sin(\varphi_j) + \sin(\xi_i)) \\ k_z = k (\cos(\varphi_j) + \cos(\xi_i)) \end{cases} \quad (26)$$

and

$$\begin{cases} k = \frac{k_x^2 + k_z^2}{2(k_x \sin(\varphi_j) + k_z \cos(\varphi_j))} \\ \xi_i = \arctan\left(\frac{2k_x k_z \cos(\varphi_j) + (k_x^2 - k_z^2) \sin(\varphi_j)}{2k_x k_z \sin(\varphi_j) + (k_z^2 - k_x^2) \cos(\varphi_j)}\right) \end{cases} \quad (27)$$

As a consequence, in the general case where the emitted field is a steered PW with angle φ_j and the received field is supposed to be a steered PW with angle ξ_i , Eq. (26)-(27) show that the temporal Fourier transform of the received signal still corresponds to a radial line of angle $\theta_i = f_{\varphi_j}(\xi_i)$ of the object spectrum (with $f_{\varphi_j}(\cdot) = \arctan((\sin(\varphi_j) + \sin(\cdot))/(\cos(\varphi_j) + \cos(\cdot)))$). Fig. 14 illustrates the k -space sampling derived from the proposed UFSB method according to the angle of the transmitted steered PW ($\varphi_j = 0$ and $\varphi_j \neq 0$), the boundary of the different spectra being discussed in Sec. 4.8. Such relations not only allow the reconstruction of one image from the emission of a single steered PW but also allow the application of compounding scheme directly in the k -space domain. Finally, the spectrum of the image to reconstruct is averaged from a set of reconstructed spectra obtained from steered PW with different angles in transmission.

4.4 UFSB : summary and practical implementation

The generic implementation of the procedure for the proposed UFSB scheme (based on steered PW insonifications) can be summarized as follows :

1. Transmit a PW with angle φ_j . The steering angle φ_j is obtained by delaying the firing time of each element of the array. The delay d is given by $d(x_i) = x_i \sin(\varphi_j)/c$, $x_i \in (-D/2, D/2)$, where x_i is the position of the array element, c is the speed of sound,

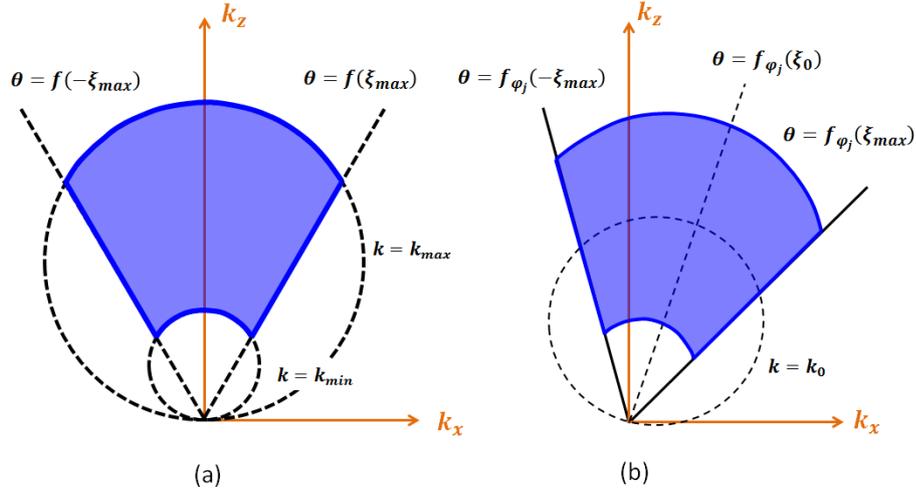


FIGURE 14 – Illustration of k -space sampling derived from the UFSB method using in emission either (a) a PW perpendicular to the probe or (b) a steered PW with direction φ_j .

and D is the aperture size. To make the system causal, an additional constant delay $dt = \left| \frac{D \sin(\varphi_j)}{2c} \right|$ is added to the delay function.

2. Receive the back-scattered echoes on each transducer element simultaneously. The 2D data set $s(x_i, t)$ is collected. Each column $s(x_i, t)$ corresponds to the received data for the transducer element at position x_i .
3. Do the 1D Fourier transform of the received data $s(x_i, t)$ over t to get the echoes in the frequency domain $S(x_i, f)$.
4. Apply delays in reception to receive steered PW with angle ξ_i . In time domain, the delays applied on the $s(x_i, t)$ should be $\tau(x_i) = x_i \sin(\xi_i)/c$, $x_i \in (-D/2, D/2)$ to receive a steered PW with angle ξ_i . Since there is a constant delay dt that has been added in transmission, it should be compensated in reception. The delay applied in reception is then $\tau - dt$. According to the shift property of the Fourier transform, a multiplication of complex exponential $\exp(-j2\pi f(\tau - dt))$ is applied on $S(x_i, f)$. One radial line of the 2D Fourier spectrum of the object is obtained by performing the summation of the corresponding signals $S(x, f) \exp(-j2\pi f(\tau - dt))$ along x_i .
5. Repeat step 4 with different angle ξ_i to fully sample the object spectrum, then interpolate the radial sampled spectrum on a regular grid.
6. Apply the 2D inverse Fourier transform to reconstruct the corresponding RF image.
7. If successive steered PW are used to perform coherent compounding, repeat steps 3 to 6 for each firing and average all the reconstructed RF images to get the final compounded image.

Note : The compounding step can also be done in the Fourier domain by averaging the obtained spectrums. This means to *i*) repeat step 3 to 5 to get the spectrums of each firing and average all the spectrums from each steered PW ; *ii*) do the 2D inverse Fourier transform on the compounded spectrum to reconstruct the final image.

4.5 Results

4.6 Numerical simulations

Numerical simulations were performed to assess the performance of the proposed UFSB PW imaging method. For all the tests, we used the dichotomous transmission scheme described in Sec. 3 with an optimal steered angle α_s of 10° . The proposed UFSB method was compared with the current state-of-the-art ultrafast PW imaging algorithms, *i.e.* the methods of Lu *et al.* [154], Garcia *et al.* [90] and Montaldo *et al.* [167]. A standard linear-array probe of 128 elements with 0.19-mm pitch, 13-mm height, a center frequency of 5-MHz (100% bandwidth), and a 23-mm elevation focus was implemented using Field II [120]. The sampling frequency was set to 100 MHz. No apodization was used in both transmission and reception. Each reconstructed image was obtained from the following protocol : *i)* raw-data signals were acquired using steered PW with different transmit angles ; *ii)* those signals were processed independently and then averaged to get the compounded RF image ; *iii)* the corresponding envelope image was derived through a Hilbert transform and normalized ; *iv)* the envelope image was then gamma-compressed using $\gamma = 0.3$ and finally converted to 8-bit grayscale to get the B-mode image (except for the lateral resolution study).

4.6.1 Validation of the UFSB method

We first used a numerical phantom composed of a discrete set of point scatterers to verify the proposed UFSB method with PW in transmission steered at different angles and for the compounding scheme. Fig. 15 displays the corresponding images reconstructed from the proposed UFSB method with different transmission angles of -10° (Fig. 15a), 0° (Fig. 15b), 10° (Fig. 15c) and the compounded image of the 3 PW ($-10^\circ, 0^\circ, 10^\circ$) in transmission (Fig. 15d), respectively. The red circles correspond to the real position of the scatterers. From the first 3 figures, it can be seen that the angle of transmitted PW has obviously a direct influence on the direction of the side lobes of the point spread function. After coherent compounding of the images from different steered PW (Fig. 15d), the side lobe artifacts can be reduced significantly, resulting in a better contrast and resolution quality.

4.6.2 Image quality evaluation

Lateral resolution

The lateral resolution of the reconstructed images was then evaluated using the numerical phantom displayed in Fig. 15. The corresponding values were measured as the width at -6 dB of the point spread function associated to the points located at 40-mm in the image. Fig. 16 investigates the influence of the number of PW on the lateral resolution using the different reconstruction techniques. First, it can be seen that the 4 methods produced similar lateral resolution whatever the number of involved steered PW. In particular, it can be observed that the lateral resolution improves quickly for 3 PW and then tends to stabilize to the optimal value, *i.e.* around 0.65-mm at 40-mm depth. It is also important to note that the lateral resolution tends to increase a little bit after the compounding of 3 PW. This behavior can be explained by the chosen dichotomous transmission scheme described in Sec. 3.2.2. From this figure, one can see that the lateral resolution measured in simulations for increasing number of PW is coherent with the theoretical evolution (Fig. 5).

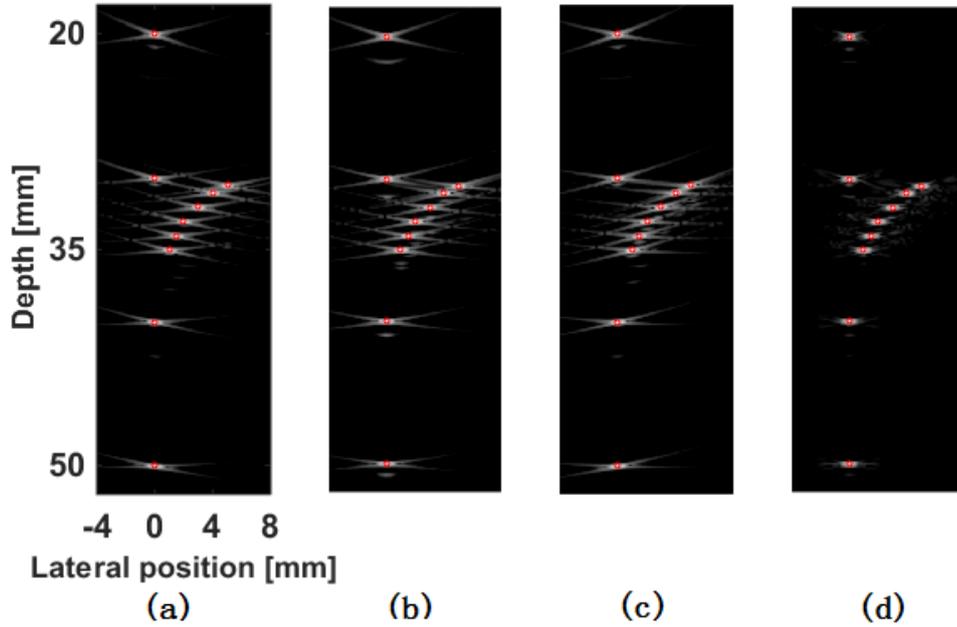


FIGURE 15 – Point scatterer images reconstructed from the UFSB method using 1 PW in transmission with steered angles of (a) -10° (b) 0° and (c) 10° . (d) Compounded image of the previous 3 steered PW. The red circles correspond to the true position of the involved scatterers.

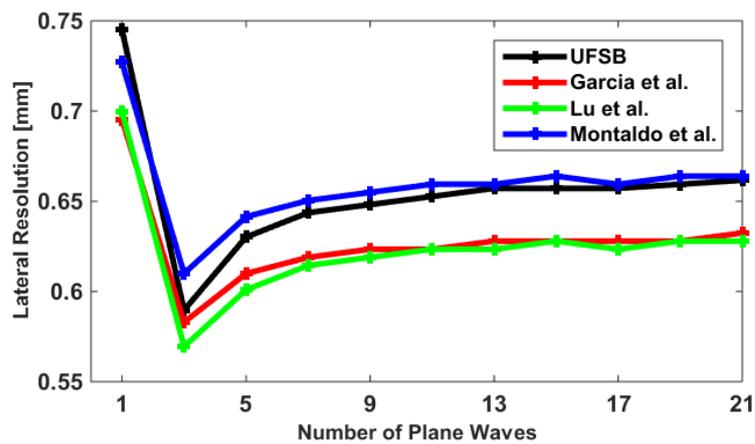


FIGURE 16 – Lateral resolution as a function of the number of PW computed at 40-mm depth.

Image contrast

The quality of the reconstructed image was then investigated from the contrast ratio (CR) measurements performed through simulated phantoms (*i.e.* a homogeneous medium with high density of scatterers, *i.e.* 20 per resolution cell, with an anechoic cyst of diameter 8-mm lying at 30-mm depth). Fig. 17 investigates the improvement of the image contrast with the number of steered PW. First, it can be seen that the 4 methods produced similar image quality both for a single PW (Fig. 17(a)) or using 21 steered PW with coherent compounding (Fig. 17(b)). These observations are confirmed by the graph provided in Fig. 17(c) which displays the evolution of the CR as a function of the number of steered PW. Interestingly, this result shows that among all the reconstructed methods, the proposed UFSB method provides almost the same CR value as the method of Lu, which further reveals that the UFSB technique provides very competitive contrast measurements compared to the state-of-the-art methods.

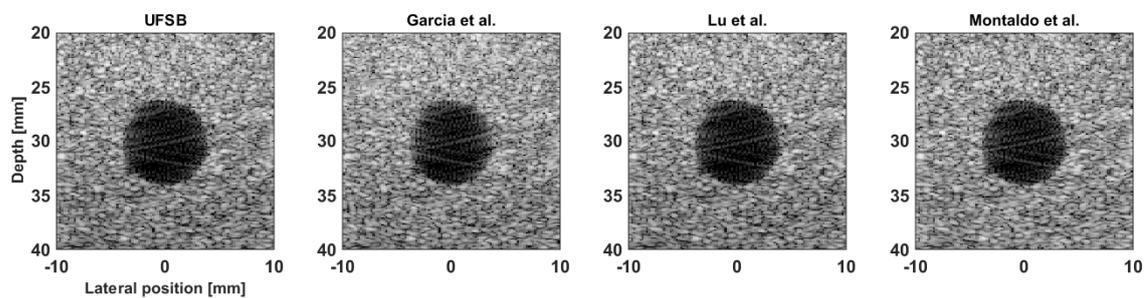
4.7 *In vitro* and *in vivo* experiments

4.7.1 *In vitro* experiments

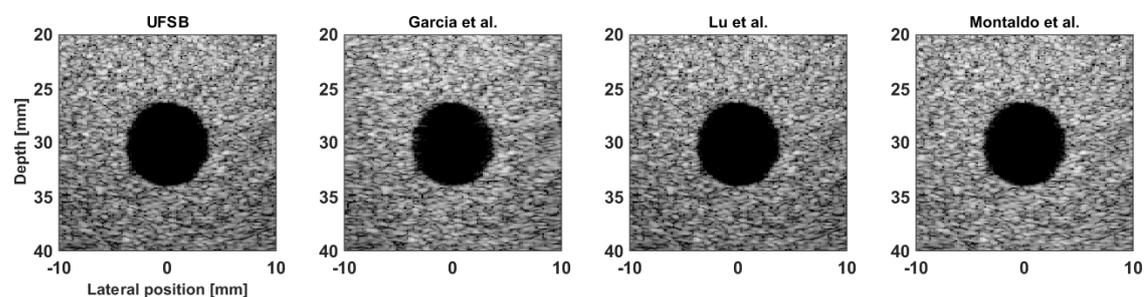
The imaging quality was also assessed experimentally using an ultrasound phantom (CIRS model : 054GS) [1]. Fig. 18 displays the schematic diagram of the corresponding phantom along with the two imaging planes obtained in this experiment. A standard linear array probe (128 elements, center frequency of 7.8-MHz, transmitted frequency of 5-MHz, Verasonics - L12-5-50mm) with the same characteristics as the ones used in simulations was interfaced with a Verasonics system to image the phantom. The imaging depth was set to 50-mm. The number of transmitted PW varied from 1 to 21, providing a frame rate between 15400 and 733 fps, respectively. No apodization in transmission and reception was used for the entire acquisitions. The received raw-data were processed using the proposed UFSB technique for PW, the Fourier-based methods of Lu *et al.* [154] and Garcia *et al.* [90] and the spatial approach of Montaldo *et al.* [167]. For all the tests, we use the dichotomous transmission scheme described in Sec. 3.2.2 with an optimal steered angle α_s of 10° .

Lateral resolution

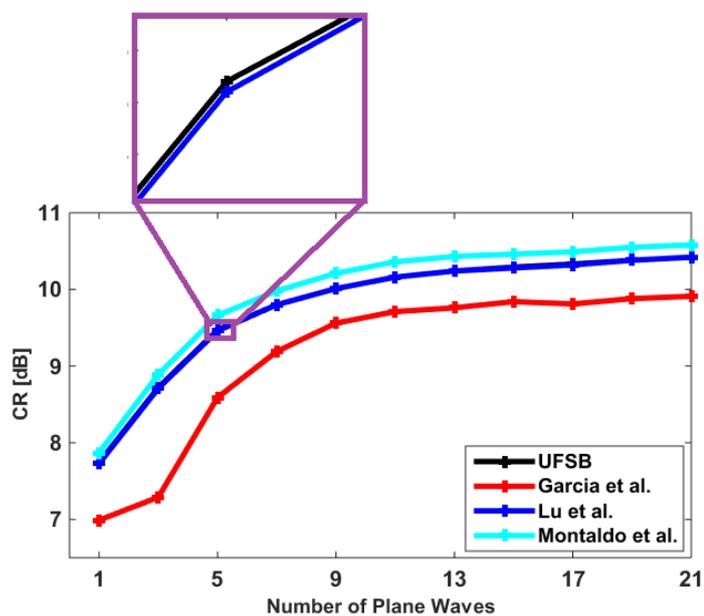
The lateral resolution was first investigated experimentally from the imaging plane n°2 given in Fig. 18 (acquisition centered on the 0.1-mm nylon monofilament targets). The corresponding values were measured as the width at -6 dB of the PSF corresponding to the two points present at 20-mm and 40-mm in the image. Fig. 19 investigates the improvement of the lateral resolution with the number of steered PW for the different reconstructed techniques and for different depths. First, it can be seen that the 4 methods still produced similar image quality both for a single PW (Fig. 19(a)) and using 21 steered PW with coherent compounding (Fig. 19(b)). These observations are confirmed by the two graphs provided in Fig. 19(c) which display the evolution of the lateral resolutions measured at 20 and 40-mm with respect to the number of steered PW. Using the dichotomous transmission scheme described in Sec. 3.2.2, it can be observed that the lateral resolution improves quickly for 3 steered PW and then tends to stabilize to the optimal value whatever the reconstructed method, *i.e.* around 0.9-mm at 20-mm depth and between 0.6 and 0.8-mm at 40-mm depth, which is in good agreement with what has been observed previously from simulations.



(a)



(b)



(c)

FIGURE 17 – B-mode images of an anechoic cyst phantom reconstructed from the four PW imaging methods using (a) 1-PW and (b) 21-PW with coherent compounding. (c) Contrast ratio CR as a function of the number of steered PW for the 4 reconstruction methods.

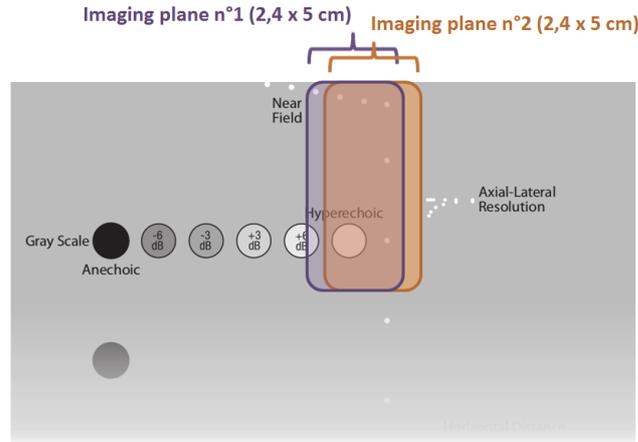


FIGURE 18 – Schematic diagram of the CIRS tissue-mimicking phantom with the corresponding imaging planes used in the experiments

Image contrast

The image contrast was then investigated from the imaging plane n°1 given in Fig. 18 (acquisition centered on the hyperechoic target). Fig. 20 shows the improvement of the image contrast with respect to the number of steered PW for the four reconstruction methods. It can be observed that the four methods yield very close contrast ratio whatever the number of steered PW. In particular, the image contrast improves significantly for a small number of steered PW and tends to stabilize for more than 9 steered PW, which is also consistent with the simulation results.

4.7.2 *In vivo* experiments on a carotid

An *in vivo* experiment was finally carried out on the carotid artery of a healthy volunteer. The transducer used in this experiment is a standard linear array probe (128 elements, center frequency of 7.8-MHz, transmitted frequency 7.8-MHz, Verasonics - L12-5-50mm) with the same characteristics as the ones used in simulation section. The imaging depth was set to 50-mm. The number of transmitted steered PW varied from 1 to 21, providing a frame rate between 15400 and 733 fps, respectively. No apodization in transmission and reception is used for the entire acquisitions. The received raw-data were processed using the three state-of-the-art techniques for PW imaging described in Sec. 2. The optimal steered angle α_s was set to 10° . The beamformed RF data were Hilbert transformed and normalized to get the envelope images. The corresponding images were then log-compressed to get the B-mode images with a 40-dB dynamic range. Fig. 21(a) and 21(b) shows the carotid images obtained with the proposed UFSB method and the state-of-the-art PW imaging methods using 1 PW in transmission and 21 steered PW in transmission with coherent compounding, respectively. As in simulations and *in vitro* experiments, these results illustrate visually the closeness of the images produced by the different methods in terms of image quality and speckle definition. In order to further investigate the influence of the four reconstruction methods, we display in Fig. 21(c) the normalized absolute differences of the carotid images given in Fig. 21(b). From this figure, it can be seen that the differences between UFSB and the other three methods are quite small. Particularly, the difference between UFSB and Lu is almost null for the imaging region (*i.e.* from 5-mm to 40-mm depth).

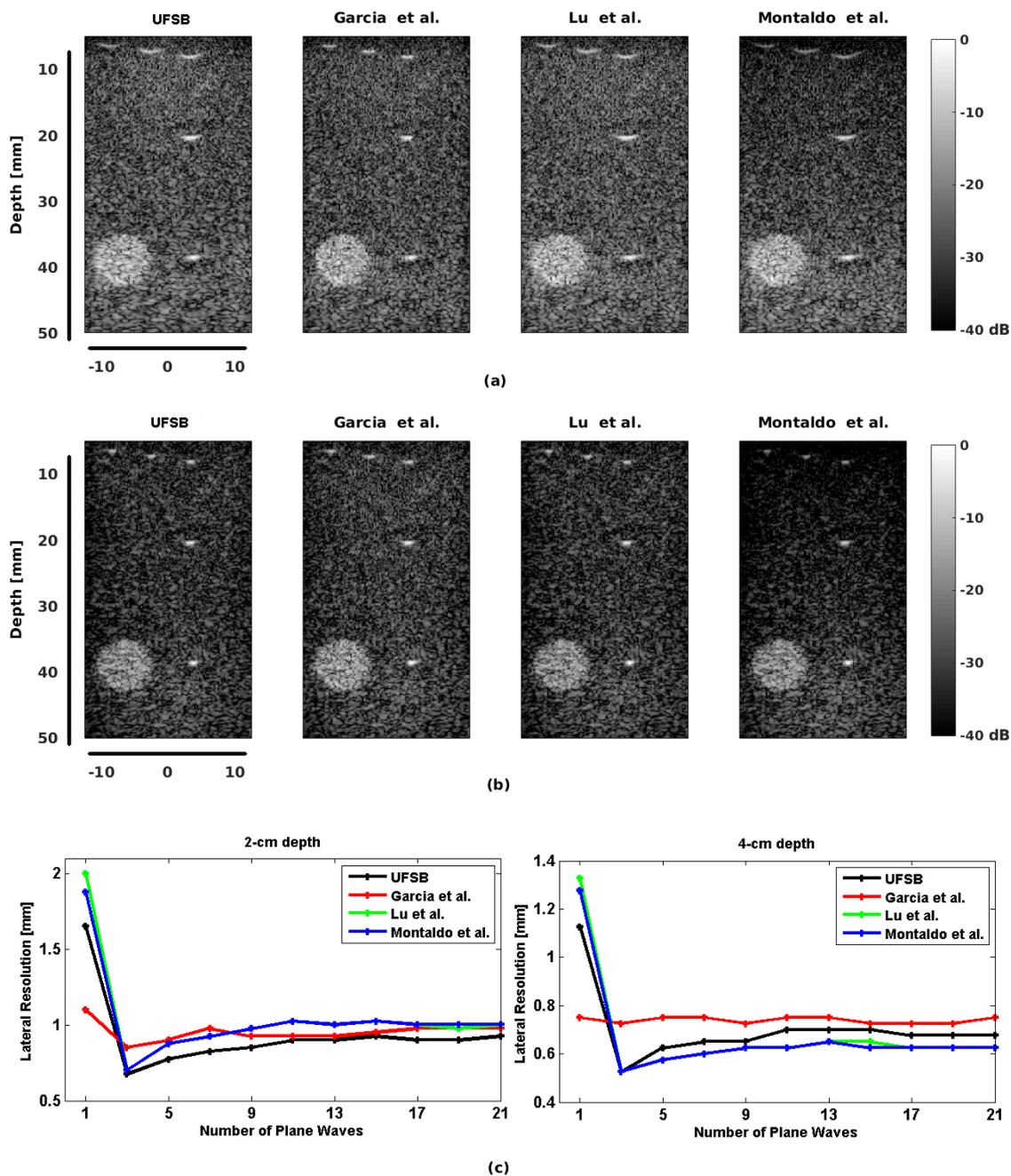


FIGURE 19 – B-mode images of a CIRS phantom reconstructed from the four reconstructed methods using (a) 1 PW and (b) 21 steered PW with coherent compounding. (c) Lateral resolution measurements as a function of the number of steered PW.

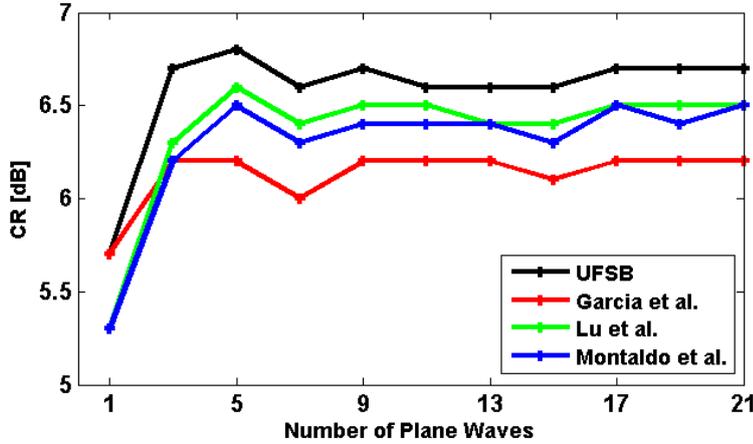


FIGURE 20 – Contrast ratio CR as a function of the number of steered PW for the four PW imaging methods.

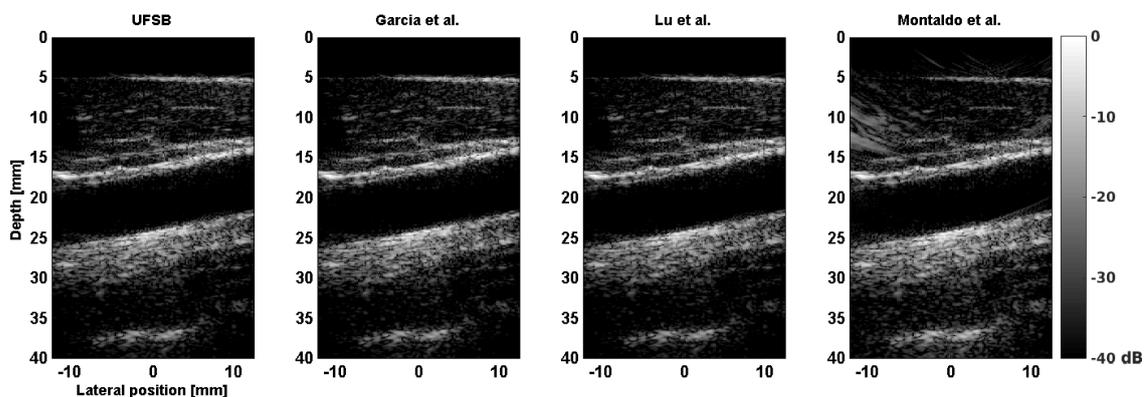
4.8 Discussion

We proposed in this section a new reconstruction scheme based on the Fourier slice theorem. Starting from the same ultrasound modeling system as in [154], we showed that ultrafast imaging in ultrasound can be represented as a problem of spectrum reconstruction based on the Fourier slice theorem. As illustrated in Fig. 13, this new technique allows recovering the Fourier spectrum of the object function radially. From the simulation and experimental results provided in Sec. 4.6 and 4.7, it can be observed that the UFSB technique produces very competitive results compared to the other existing approaches, the closest method in terms of image quality being the one proposed by Lu *et al.*. It is thus of interest to evaluate the proposed UFSB framework and the two others Fourier-based approaches through the comparison of the properties of the Fourier spectrum reconstructed from each method.

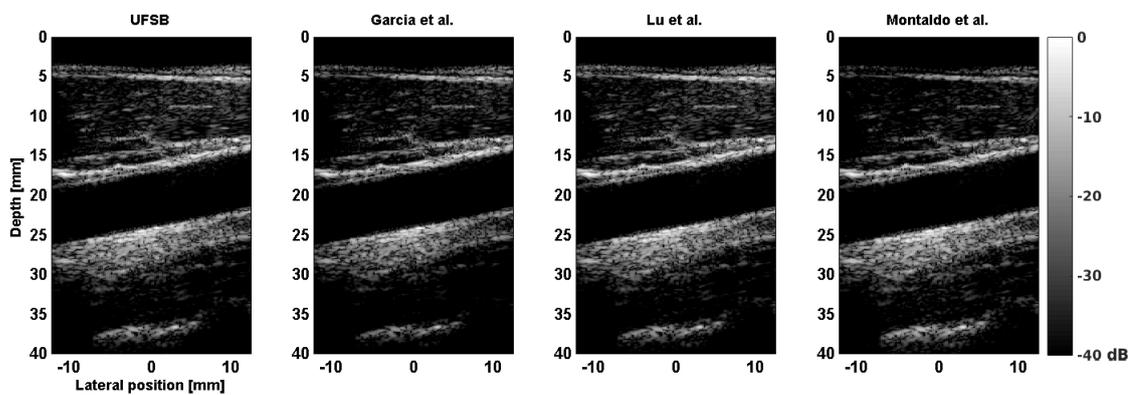
The spectral content of the object function is determined by the physical limits of the ultrasound system. In our experiments, the probe consisted of transducer elements with 100% bandwidth. This means that the spectral content along the k_z axis ($k_z = 2f/c$) is defined between $\{f_0/c, 3f_0/c\}$, where f_0 is the center frequency of the transducer. The lateral sampling frequency of the object function is conditioned by the pitch p of the probe, limiting the spectral content along the k_x axis to the range $\{-1/(2p), 1/(2p)\}$. As detailed in Sec. 4.2.1, the UFSB technique allows sampling the Fourier spectrum of the object function radially with angle $\theta_i = f_{\varphi_j}(\xi_i)$, where φ_j and ξ_i correspond to the angle of the steered PW in emission and reception, respectively. By playing with different delay strategies applied on the received signals, it is thus possible, for only one emitted PW with angle φ_j , to receive steered PW with different angles and thus radially recover the Fourier space of the object. Because of the physical limits described above, we constrained for each experiment the value of θ_i in the interval $-\theta_{max} \leq \theta_i \leq \theta_{max}$, where θ_{max} is defined as

$$\theta_{max} = \arctan\left(\frac{k_{x,max}}{k_{z,min}}\right) = \arctan\left(\frac{c}{2pf_0}\right) \quad (28)$$

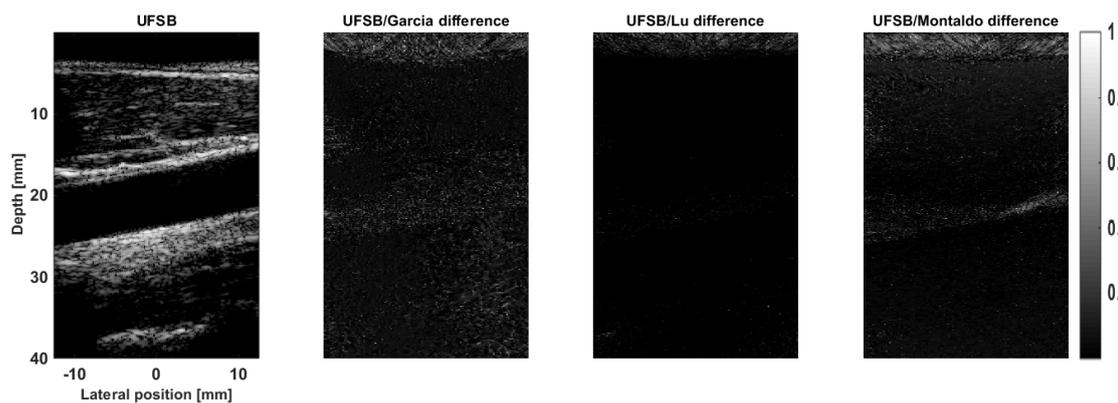
From Eq. (24) given in Sec. 4.2.1, it can be observed that for a fixed value of ξ_i , the UFSB k-space sampling corresponds to a line passing through the origin with an angle $\theta_i = f(\xi_i)$. Therefore the accessible range for θ_i corresponds to a cone of apex $2\theta_{max}$. In the same way, for a



(a)



(b)



(c)

FIGURE 21 – *In vivo* experiments - Carotid image reconstructed from the four PW imaging methods using (a) 1 PW and (b) 21 steered PW with coherent compounding. (c) Absolute differences of the UFSB and the other 3 reconstructed methods from the figures in (b).

fixed value of k , the UFSB k-space sampling corresponds to a circle centered at $(k_x, k_z) = (0, k)$, with a diameter of $2k$. Since the two conditions must be verified simultaneously, the generic k-space sampling derived from the proposed UFSB technique corresponds to the diagrams provided in Fig. 14

Contrary to the proposed UFSB scheme, the two others Fourier-based techniques proposed by Lu *et al.* [154] and Garcia *et al.* [90] sample the Fourier spectrum of the object function along the k_z axis direction with a lateral step proportional to the inverse of the pitch. In order to investigate the influence of the different Fourier sampling schemes, we display in Fig. 22 the normalized absolute differences of the Fourier spectrum of the experimental images provided at the top row of Fig. 19 reconstructed from the three different Fourier techniques. From this figure, it may be first seen that the difference of spectrum between Lu and Garcia methods is null for the central part of the spectrum, as it has been already observed in [90]. Moreover, this figure also reveals that the Fourier spectrum of the proposed UFSB technique is close to the one obtained from Lu *et al.*, the difference being uniformly distributed over the Fourier space with a relative difference of the order of 15% of the normalized values. Even if the Fourier sampling scheme is different, the proposed UFSB technique and the one of Lu *et al.* reconstruct close Fourier spectrum, the differences arising from the interpolation between the samples which are not distributed in the same manner. This result reinforces the experimental observations that reconstructions provided by the two methods are close and the differences between z-axis sampling and radial sampling are marginal.

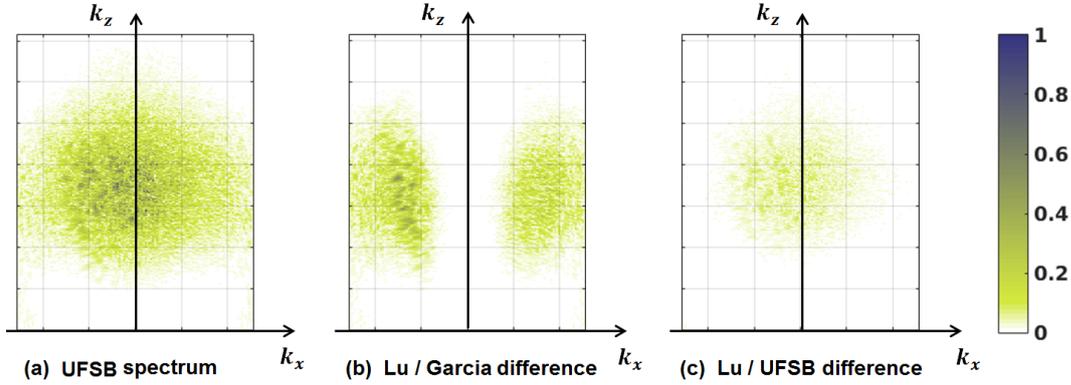


FIGURE 22 – Spectrum mapping obtained from the experimental images given at the top row of Fig. 19. (a) Magnitude of the spectrum reconstructed using the proposed UFSB technique. (b) Absolute differences of Lu’s and Garcia’s spectrum. (c) Absolute differences of Lu’s and UFSB’s spectrum.

Another strong interest of performing PW image reconstruction in Fourier domain concerns its lower computational complexity. As introduced in [90], the computational complexity decreases from $O(n_e n_e n_s)$ for the DAS technique to $O(n_e n_s \log(n_e n_s))$ for the method of Lu and Garcia thanks to Fourier formulation, where n_e denotes the number of elements in the transducer and n_s denotes the number of time samples. The reconstruction of one image using the UFSB method requires n_e 1-D FFTs (complexity of $O(n_e n_s \log(n_s))$), one 2D FFT (complexity of $O(n_x n_s \log(n_x n_s))$), the computation of $(n_x n_e n_s / 2)$ multiplications for the delays applied in reception (the ratio of 2 being explained by the use of half of the spectrum) and two times the computation of $(n_x n_s)$ interpolated values (one for the regridding step performed in the Fourier

space and one for the spatial transformation introduced in this work), where n_ξ represents the number of angles used in reception. The computational complexity of the proposed UFSB method thus corresponds to $O(n_\xi n_e n_s)$, making this method less attractive than the other two Fourier-based methods. Nethertheless, the proposed UFSB method provides stimulating conceptual link between ultrasound image formation and other imaging modalities such as CT.

4.9 Conclusions

In this section, we developed an alternative Fourier-based ultrafast imaging method named Ultrasound Fourier Slice Beamforming (UFSB). The theoretical aspect of the UFSB framework is based on the exploitation of the standard Fourier slice theorem to reconstruct ultrasound images. The reconstructive procedure, the spectrum boundary and the spectrum comparison with two others Fourier-based methods were investigated through simulations and experiments. Results revealed that the UFSB technique produces very competitive results in terms of image quality compared to the state-of-the-art methods but with either equivalent or higher computational complexity, making our solution less attractive.

5 2nd contribution : extension of Fourier-based techniques for ultrafast imaging with diverging waves

In this section we present a novel method dedicated to an efficient reconstruction of ultrasound sectorial images using DW. Our formalism is based on the derivation of an isomorphism in terms of travel time between a planar system based on PW and a sectorial one based on DW. In particular, we proposed an explicit spatial transformation, which allows the reconstruction of wide-angle images from the compounding scheme in transmission, the optimal quality of images being reached for 15 DWs (frame rate of 855 frames/s) on both numerical and in vitro phantoms. This work has been the subject of several publications [294–298].

5.1 Motivations

Thanks to the literature review proposed in Sec. 2, it appears that the different ultrafast imaging methods based on PW or DW can be classified into two groups : the Fourier-based techniques where the received raw-data are used to reconstruct the Fourier spectrum of the image of interest [90, 154] and the spatial-based approaches where the images are directly computed from the space domain [101, 167, 191]. The main advantage of the Fourier-based approaches concerns the computational complexity which has been investigated in several studies in the context of planar acquisition from steered PW [90, 132]. In particular, the computational complexity of the DAS method proposed in [167] is equal to $O(n_e n_e n_s)$ and decreases to $O(n_e n_s \log(n_e n_s))$ for Fourier-based techniques [90] (n_e representing the number of transducer elements and n_s the number of samples of the beamformed signal). Despite their computational interest, Fourier-based techniques have not been proposed so far to reconstruct sectorial images from DW. Based on this observation, we developed in this study an explicit transformation which allows a direct extension of existing Fourier-based techniques derived for planar acquisition to the reconstruction of sectorial images with wide angles (90°). The key concept of the proposed formulation is based on the derivation of an isomorphism in terms of travel time between a planar system based on PW and a sectorial one based on DW.

5.2 Methodology

In this section, we propose a formulation which allows a direct extension of the Fourier-based methods of Lu *et al.* [154] and Bernard *et al.* [35] to sectorial acquisitions using a combination of DWs in transmission. These two methods are well adapted to the formulation given below since the underlying image formation explicitly involves the same transducer for both transmission and reception. Contrary to these approaches, the method of Garcia *et al.* is based on the exploding reflector model, leading to approximations which are not adapted in this work (see [90] for more details).

5.2.1 Extension of Fourier-based techniques to sectorial imaging

The key idea of this part is to establish an isomorphism in terms of travel time when either a DW or a PW is used in transmission as illustrated in Fig. 23. Indeed, such a relation would allow the reconstruction of a sectorial image from a DW using the same formalism as the one derived for planar acquisition. Isomorphism between travel times has been first proposed by Garcia *et al.* to adapt the seismic Stolt's migration technique to ultrasound plane wave imaging [90]. Let us define a probe located at $z = 0$, centered at $x = 0$ and where each transducer element E is

positioned at $(x_E, 0)$. Getting inspired by the recent works performed on ultrafast imaging for sectorial acquisition [101, 191, 203], we propose in this work to use DW in order to insonify the entire field of view with one transmission. Such a wave is obtained through the excitation of a virtual source located behind the probe and placed at (x_v, z_v) , where $x_v \in [-L/2, L/2]$ (L being the width of the probe), and $z_v \leq 0$ is determined by the chosen angular aperture α and the sub-aperture width D as :

$$z_v = -\frac{D/2}{\tan(\alpha)} \quad (29)$$

The transmit delays Δ_E associated with the virtual source position and computed for each transducer element are then calculated as :

$$\Delta_E = (\sqrt{(x_E - x_v)^2 + z_v^2} + z_v) / c \quad (30)$$

where the second term in the brackets is used as an offset to ensure that the smallest transmitted delay is null. In this context, the travel time of a DW to reach a point scatterer positioned at (x_d, z_d) in the medium and to come back to a transducer element placed at $(x_E, 0)$ is given by :

$$\tau_d(x_E) = \left(\sqrt{(x_d - x_v)^2 + (z_d - z_v)^2} + z_v + \sqrt{(x_d - x_E)^2 + z_d^2} \right) / c \quad (31)$$

In the case of PW with normal incidence, the travel time involved to reach a point scatterer (x_p, z_p) and to come back to the same transducer element $(x_E, 0)$ is given by :

$$\tau_p(x_E) = (z_p + \sqrt{(x_p - x_E)^2 + z_p^2}) / c \quad (32)$$

By equating the first order Taylor approximation of Eq. (31) and Eq. (32) at $x_E = x_v$, the following relation between (x_d, z_d) and (x_p, z_p) can be derived :

$$\left\{ \begin{array}{l} x_p \approx \frac{(x_d - x_v) (\sqrt{(x_d - x_v)^2 + (z_d - z_v)^2} + z_v)}{z_d + \sqrt{(x_d - x_v)^2 + z_d^2}} \\ \quad + \frac{(x_d - x_v) \sqrt{(x_d - x_v)^2 + z_d^2}}{z_d + \sqrt{(x_d - x_v)^2 + z_d^2}} + x_v \\ z_p \approx \frac{z_d (\sqrt{(x_d - x_v)^2 + (z_d - z_v)^2} + z_v + \sqrt{(x_d - x_v)^2 + z_d^2})}{z_d + \sqrt{(x_d - x_v)^2 + z_d^2}} \end{array} \right. \quad (33)$$

and

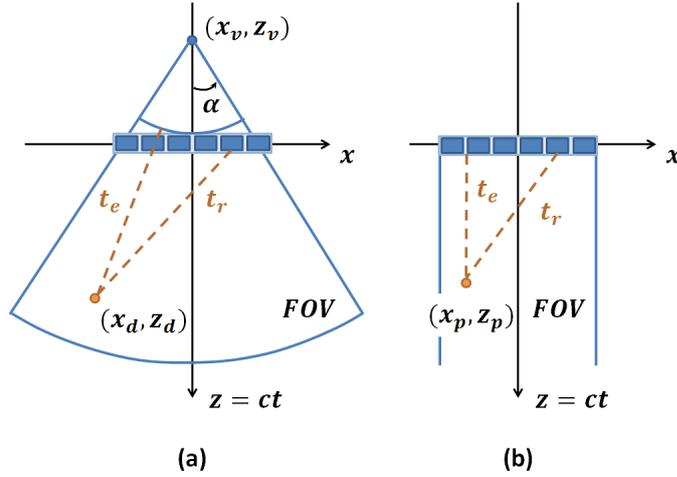


FIGURE 23 – Illustration of the travel time (t_e and t_r) involved when either (a) a DW or (b) PW is used in transmission. In both cases, ($t_e + t_r$) corresponds to the time needed by the insonifying wave to reach a scatterer (orange dot) and then come back to a specific transducer element.

$$\left\{ \begin{array}{l} x_d \approx \frac{1}{2} \frac{(x_p - x_v) \left(z_v^2 - (z_p + \sqrt{(x_p - x_v)^2 + z_p^2} - z_v)^2 \right)}{z_v z_p - (z_p + \sqrt{(x_p - x_v)^2 + z_p^2} - z_v) (\sqrt{(x_p - x_v)^2 + z_p^2})} \\ \quad + x_v \\ z_d \approx \frac{1}{2} \frac{z_p \left(z_v^2 - (z_p + \sqrt{(x_p - x_v)^2 + z_p^2} - z_v)^2 \right)}{z_v z_p - (z_p + \sqrt{(x_p - x_v)^2 + z_p^2} - z_v) (\sqrt{(x_p - x_v)^2 + z_p^2})} \end{array} \right. \quad (34)$$

From Eq. (33) and Eq. (34) the following algorithm is thus proposed to reconstruct a sectorial image : *i*) transmission of a DW using a standard phased-array ; *ii*) from the received echo, application of a standard Fourier-based technique derived for PW. From Eq. (33), one can see that the output of this step will reconstruct the desired ultrasound image but expressed in the (x_p, z_p) coordinate system ; *iii*) application of the spatial transformation given through Eq. (34) which allows expressing the reconstructed image in the conventional cartesian space.

5.2.2 Correctness of the proposed formulation

It is worth pointing out that Eq. (33) and Eq. (34) have been obtained for the particular case $x_E = x_v$, meaning that the equivalence between the travel time of Eq. (31) and Eq. (32) is exact only near the region defined by $x = x_v$, where the Taylor approximation remains valid. We thus investigated the error in terms of travel time difference for the special case $x_v = 0$ when the relation provided by Eq. (33) and Eq. (34) are used for each point in the medium. More precisely, for each point (x_d, z_d) of the medium, we first computed the set of travel times $\{\tau_d(i)\}_{i \in [1, N]}$ from the virtual source point $(0, z_v)$ to (x_d, z_d) back to each transducer element i of the probe (N being the number of elements). We then derived the corresponding (x_p, z_p) points and compute the set of travel times $\{\tau_p(i)\}_{i \in [1, N]}$ from the probe to (x_p, z_p) back to each

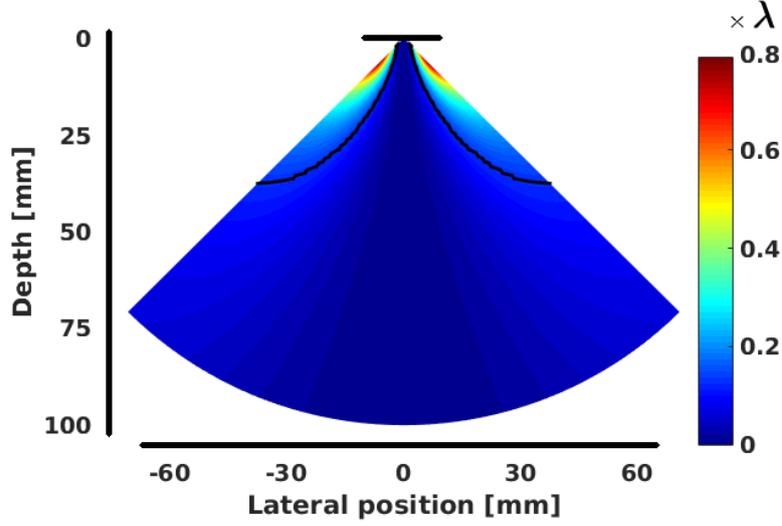


FIGURE 24 – Error map derived from the travel time equivalence between a sectorial acquisition using DW and the equivalent PW imaging after the application of the proposed spatial transformation. The computed error is expressed relatively to the wavelength used in the experiments. The black line on the top of the figure models the transducer location while the axis origin is located at the center of the probe. The two black curves drawn inside the error map represent the boundary of the region at the edge of the image where the maximum error is higher than $\lambda/8$

transducer element i . In the case of a perfect equivalence between the two systems, the set of the pair-wise difference $\tau_d(i) - \tau_p(i)$ over the probe element should be zero. Fig. 24 shows the corresponding maximum value (*i.e.* $\max_i(\tau_d(i) - \tau_p(i))$) computed for each point of the medium in terms of the wavelength λ . From this figure, one may first observe that there is no difference in terms of travel time at the center of the probe and this error increases at the edge of the image. The maximum error appears in the near field (depth lower than 1-cm) at the edges of the image with a value around $0.8 \times \lambda$. This figure also allows the assessment of the potential defocusing effect induced by the proposed formulation. Indeed, each value displayed in Fig. 24 corresponds to the maximum of the error between the two systems computed over the full aperture. Since for each reconstructed point, the travel time error for most of the elements is below $\lambda/8$ (which is the commonly accepted value under which errors in travel distances can be neglected), it reveals the marginal effect of defocusing during the beamforming process. Finally, it is worth pointing out that the neglected effect of defocusing also justifies the correct reconstruction of the point positioning after the spatial transformation. This aspect will be further investigated in Sec. 5.4.

5.2.3 Summary and practical implementation

The formulations introduced in Sec. 5.2.1 allow the reconstruction of an image from the transmission of a diverging wave with a virtual source positioned at any (x_v, z_v) . Thus, by reproducing the scheme proposed in [191], *i.e.* emitting several diverging waves with virtual sources positioned at $z_v = -D/(2 \tan(\alpha))$ and for different x_v values, our formalism can be used to reconstruct wide field of view ultrasound images with efficient compounding scheme. The implementation of the procedure for sectorial imaging from the two Fourier-based techniques described in Sec. ?? and Sec. 4.2 can thus be summarized as follows :

1. Transmit one or several DWs to insonify the medium. In particular, for a virtual source positioned at (x_v, z_v) , the corresponding DW is generated by applying the delays given in Eq. (4) for each transducer element.
2. Receive simultaneously on each transducer the backscattered echoes. This step allows the collection of a 2D matrix for each DW insonification.
3. Apply one particular Fourier-based reconstruction method designed for PW (i.e. Lu or Bernard's method) on each received 2D matrix. This step allows the reconstruction of an RF image which is expressed in a space with coordinates (x_p, z_p) .
4. Apply the spatial transformation given in Eq. (8) to reconstruct the corresponding RF image in the Cartesian space with coordinates (x_d, z_d) .
5. If successive DWs are used to perform coherent compounding, repeat steps 3 and 4 for each firing and average all the reconstructed RF images to get the final compounded image.

Note that step 3) implies the interpolation of the collected data on a regular grid in the Fourier domain in order to compute the reconstructed image using a regular 2D inverse Fourier transform. Such interpolation is an old topic which has been widely studied in the literature [112, 188, 248]. In particular, as described in details in [188], the linear interpolation applied in k-space leads to the multiplication of the real image by a squared sinc function. This results in a decrease of the intensity of the current reconstructed image along depth. To reduce this artifact, we applied a depth-varying intensity correction for the Fourier-based techniques. An intensity correction of 0.5-dB/MHz/cm was used.

5.3 Experiments

The performance of the proposed transformation associated with the Fourier-based techniques proposed by Bernard and Lu was evaluated from both numerical and *in vitro* phantoms, as well as from *in vivo* data from a cardiac acquisition.

5.3.1 Acquisition protocol

The same standard phased-array probe of 64 elements with a center frequency of 2.5-MHz (100% bandwidth), 0.32-mm pitch, 13-mm height and 60-mm elevation focus was used for all the tests. No apodization was used in transmission or reception. The imaging depth was set to 100-mm. For all the tests, the two extended Fourier methods were compared with the current state-of-the-art spatial-based method of Papadacci *et al.* Each reconstructed image was obtained from the following protocol : *i)* raw-data signals were acquired using DW with different virtual source point positions ; *ii)* those signals were processed either following the two Fourier based methods described above or using the method of Papadacci *et al.* to obtain the compounded RF image ; *iii)* the corresponding envelope image was derived through a Hilbert transform and normalized ; *iv)* the envelope image was then gamma-compressed using $\gamma = 0.3$ as in [90] and finally converted to 8-bit grayscale to get the B-mode image.

5.3.2 Transmission scheme

For comparison purposes and because of its efficiency, we used the same transmission scheme as the one proposed in [191]. Each DW involved in transmission was emitted from a virtual source point with an angular aperture of 90° and a sub-aperture width composed of 21 elements.

For a single transmission, the virtual source was placed at the center of the probe. For three transmissions, the two others sources were placed near the edges of the transducer (*i.e.* at $x_v = [-6.7, 0, 6.7]$ -mm). For more than three transmissions, the additional virtual sources were uniformly distributed (*i.e.* at $x_v = [-6.7, -3.35, 0, 3.35, 6.7]$ -mm for 5 transmissions in our experiments). The number of transmitted DWs varied from 1 to 25, providing a frame rate between 7700 and 308 frame per second (fps).

5.3.3 Evaluation metrics

The image quality was assessed from the lateral resolution and the image contrast metrics. The contrast was measured from the B-mode images using the following classical contrast ratio (CR) [284] :

$$CR = 20 \log_{10} \frac{|\mu_t - \mu_b|}{\sqrt{(\sigma_t^2 + \sigma_b^2)/2}}, \quad (35)$$

where μ_t and μ_b (σ_t^2 and σ_b^2) are the means (variances) of gray levels in the targets and the surrounding background.

5.3.4 Numerical simulations

The conceptual correctness of the proposed spatial transformation was first investigated from a simple synthetic phantom composed of a discrete set of point scatterers (Fig. 26). The same phantom was also used to evaluate the quality of the reconstructed images through the lateral resolution. The corresponding values were measured as the full width at half maximum of the point spread function associated to the points located at 20-mm, 40-mm, 60-mm and 80-mm in the image. The image contrast was investigated from another phantom composed of a medium with high density of scatterers (20 per resolution cell) with two anechoic cysts of diameter 8-mm lying at 40 and 80-mm depth, respectively. For all the numerical simulations, the corresponding raw-data were generated using Field II [116, 118].

5.3.5 In vitro experiments

The in vitro experiments were performed with a Verasonics research scanner (V-256, Verasonics Inc., Redmond, WA) and a 2.5 MHz phased-array transducer (ATL P4-2, 64 elements). The CIRS tissue-mimicking phantom displayed in Fig. 25 was used for the in vitro experiments. The lateral resolution was first investigated from the acquisition centered on the 0.1-mm nylon monofilament targets positioned at 20-mm, 40-mm, 60-mm and 80-mm. The performance in terms of contrast was then measured from the hyperechoic cyst presents at 40-mm depth.

5.3.6 In vivo experiments

The in vivo experiments were performed with a Verasonics research scanner (V-1-128, Verasonics Inc., Redmond, WA) and a 2.5 MHz phased-array transducer (ATL P4-2, 64 elements). Very recently, Porée *et al.* have introduced a scheme for coherent compounding of tilted DWs which has been validated in cardiac imaging [203]. In particular, to deal with high myocardium velocities, the authors have proposed to use a triangle transmit sequence of DWs, combined with tissue Doppler imaging to perform motion compensation (MoCo). In order to assess the ability of our approach to deal with in vivo data, we applied the exact same transmission scheme as the one described in [203] to reconstruct a cardiac sequence with and without motion compensation. In

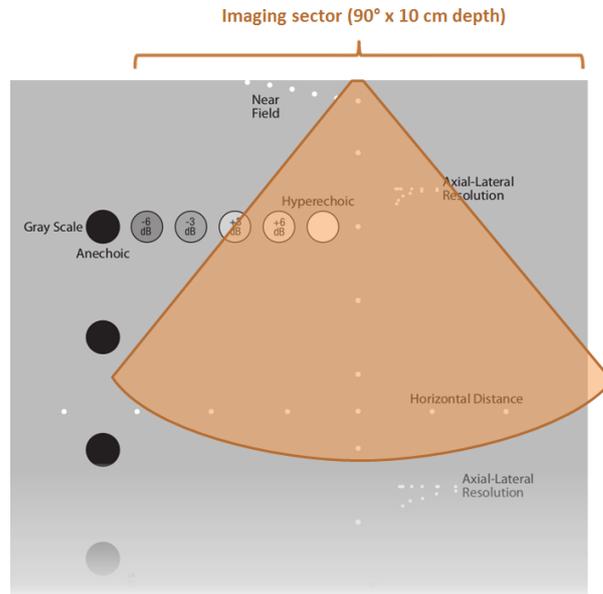


FIGURE 25 – Schematic diagram of the CIRS tissue-mimicking phantom (model : 054GS) with the corresponding imaging sector used in the DWs experiments.

particular, we used 32 DWs in transmission (tilted from -25° to $+25^\circ$ with a triangle strategy) for the reconstruction of each frame of the final sequence. The cardiac data from Porée *et al.* [203] were reanalyzed as follows : *i*) apply the triangle transmit sequence of DWs given in [203], *ii*) for each firing, reconstruct a beamformed sectorial image using the extension of Lu’s method ; *iii*) apply the MoCo algorithm described in [203] on each of the image obtained at the previous step ; *iv*) perform the compounding on the corrected images ; *v*) repeat this process for each frame of the sequence.

5.4 Results

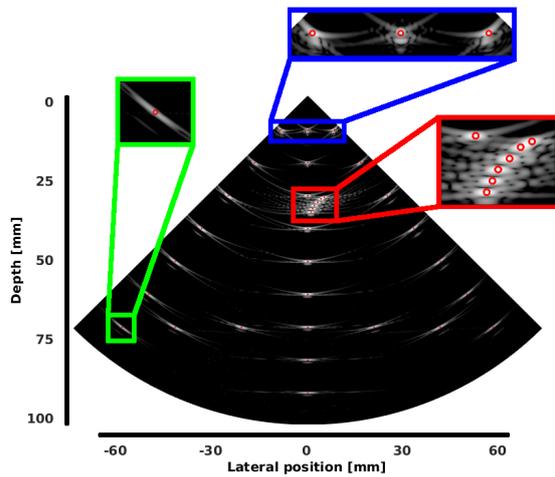
5.4.1 Numerical Simulations

Validation of the proposed framework

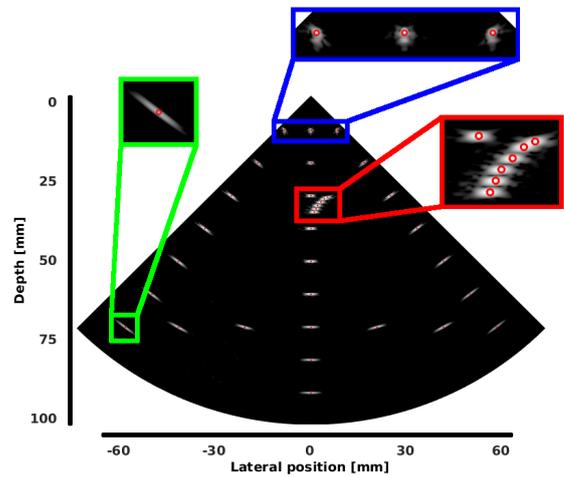
Fig. 26 displays the images reconstructed from the algorithm described in Sec. 5.2.3 associated with the Fourier-based method of Bernard (Fig. 26(a) and 26(b)) and Lu (Fig. 26(c)) as well as the image reconstructed by the method of Papadacci *et al.* (Fig. 26(d)). The red circles correspond to the real position of the scatterers. From these figures, it can be seen that the proposed transformation allows the reconstruction of all the scatterers with the correct position whatever the chosen Fourier-based technique. This illustrates the marginal effect of the error map presented in Sec. 5.2.2 and validates the flexibility of the proposed approach.

Lateral resolution

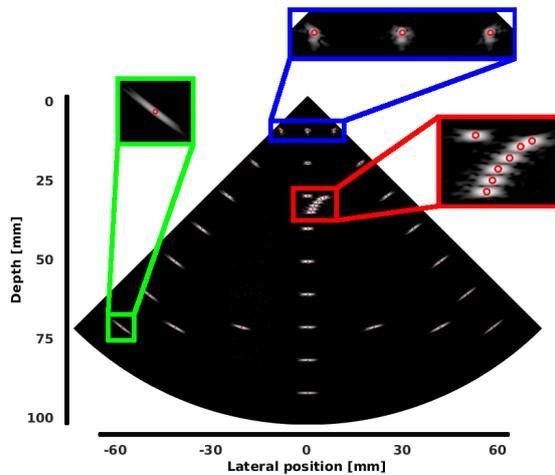
Fig. 27(a) investigates the influence of the number of DWs on the lateral resolution using the different reconstruction techniques and for the different depths at the middle of the image (azimuth angle of 0°). First, it can be seen that the 3 methods produce similar image quality whatever the depth and the number of DWs. In particular, it can be observed that the lateral resolution improves quickly for 3 DWs and then tends to stabilize to the optimal value, *i.e.* around 0.8-mm



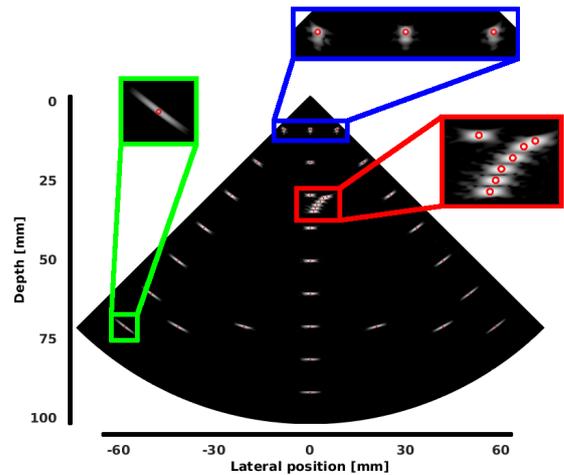
(a) Sectorial extension of Bernard's method using 1 DW



(b) Sectorial extension of Bernard's method using 15 DWs



(c) Sectorial extension of Lu's method using 15 DWs



(d) Papadacci's method using 15 DWs

FIGURE 26 – Synthetic phantom reconstructed from the extension of the Fourier-based methods of Bernard *et al.* and Lu *et al.* as well as the method of Papadacci *et al.* The red points correspond to the real position of the scatterers involved in the simulation.

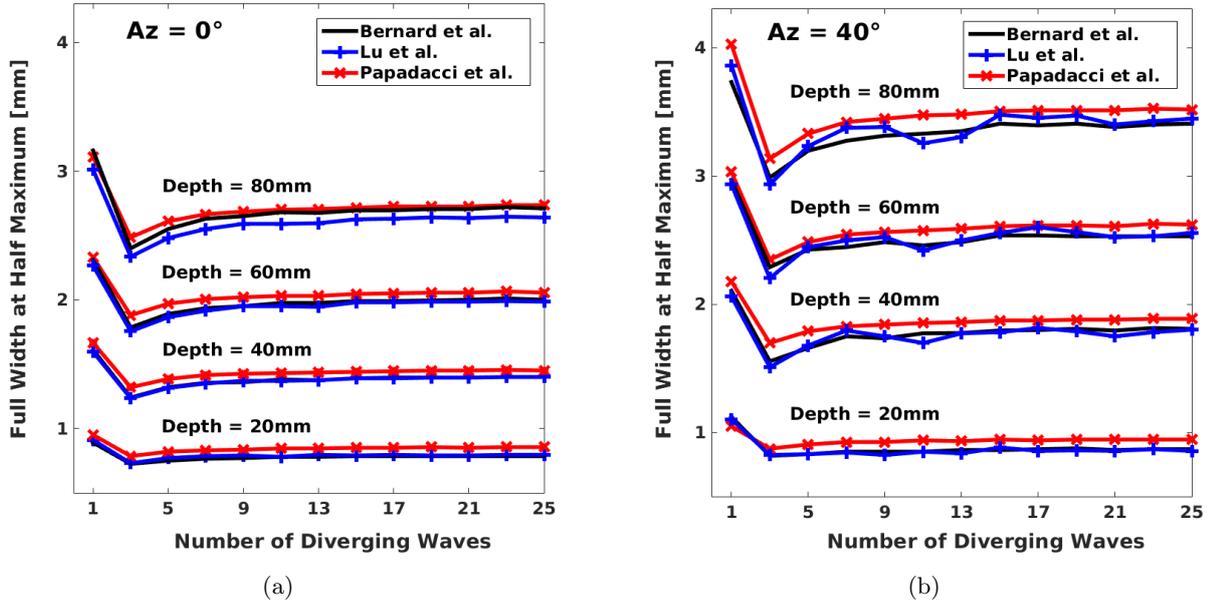


FIGURE 27 – Numerical phantom - lateral resolution measurements as a function of the number of DWs computed at 20-mm, 40-mm, 60-mm and 80-mm at (a) the middle of the image (azimuth angle of 0°) and (b) at the edge of the image (azimuth angle of 40°).

at 20-mm depth, 1.4-mm at 40-mm depth, 2.0-mm at 60-mm depth and 2.7-mm at 80-mm depth. It is also important to note that, for each depth, the lateral resolution tends to degrade a little bit after the compounding of 3 DWs. This behavior can be explained by the chosen transmission scheme. Indeed, in the particular case of a monochromatic far-field approximation at a focal distance F , the width W of the main lobe corresponding to a coherent summation of N DWs ($N > 1$) is in the order of (see Eq. (4) in [191]) :

$$W = 2 \frac{\lambda F}{W_{va}} \frac{N - 1}{N} \quad (36)$$

where $\lambda = 2\pi/k$ and W_{va} corresponds to the width of the virtual array which is considered as fixed in this study. Based on this equation, we display in Fig. 28 the evolution of the approximation of the full width at half-maximum from the probe settings used in our simulations at a focal distance of 60-mm. From this figure, one can see that the theoretical evolution of this measure is consistent with the lateral resolution measured in simulations.

Fig. 27(b) investigates the influence of the number of DW on the lateral resolution using the different reconstruction techniques and for different depths at the edge of the image (azimuth angle of 40°). As for the middle case, it can be seen that the 3 methods produce similar image quality whatever the depth and the number of the involved DWs. It is also interesting to note that the lateral resolution (computed along the azimuth direction for consistency) degrades with the azimuth angle. For instance, at 80-mm depth, the lateral resolution goes from values around 2.7-mm at the middle of the image to values around 3.4-mm at the edge. This phenomenon can be explained by the fact that the region where the different DWs contribute to the image pixels is optimal at the middle of the image, leading to a better image resolution. The limited opening angle of the effective aperture may also reinforce this phenomenon. Finally, it can be observed

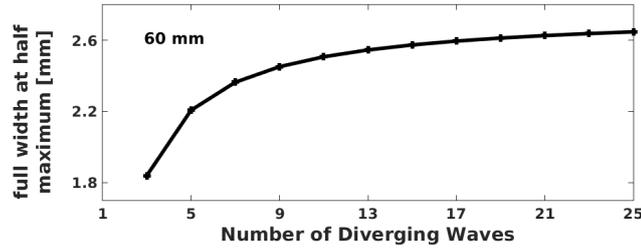


FIGURE 28 – Approximation of the full width at half maximum of the main lobe corresponding to the coherent summation of DWs in the case of a monochromatic far-field at a focal distance of 60-mm.

from Fig. 27(a) and 27(b) that the difference in terms of lateral resolution between Papadacci’s method and the Fourier based methods varies in average from 0.06-mm (at 20-mm depth) to 0.10-mm (at 80-mm), which corresponds of a maximum difference lower than $\lambda/6$. Since this difference is much lower than half of the wavelength, this phenomenon may be considered as negligible.

Image contrast

Fig. 29(a) displays the CR values measured for two different depths at the middle of the image (azimuth angle of 0°) for each compounding experiment. The three methods yield very close results, the CR measurements showing same tendency with an increasing number of DW and with better CR scores at 40-mm. This can be easily explained by the intrinsic decrease of the transmit ultrasound pressure with depth which inevitably induces lower contrast. It is also interesting to note that for more than 15 DWs, the image contrast tends to stabilize to the optimal value for all methods, which is consistent with the results provided in [191]. Finally, Fig. 29(b) to 29(d) display the images reconstructed with the three different methods using 15 DWs. These results illustrate visually the closeness in terms of image quality and speckle definition of the images reconstructed from the proposed extension of the Fourier-based techniques and the spatial-based method of Papadacci *et al.*. Fig. 30 displays the CR values measured for two different depths at the edge of the image (azimuth angle of 40°). As for the cysts in the middle, it can be seen that the 3 methods yield very similar image quality whatever the depth and the number of involved DW. It is of importance to note that the image contrast decreases a little bit with the azimuth angle. For instance, at 40-mm depth, the CR decreases from values around 13-dB at the middle of the image down to values around 10-dB at the edge.

5.4.2 In vitro and in vivo experiments

Lateral resolution

Fig. 31 illustrates the impact of DW compounding number on the lateral resolution for the different reconstruction techniques and for the different depths. All the experimental results are consistent with what have been previously observed from simulations, both in terms of tendency and measured values. Indeed, for all the methods, the lateral resolution improved quickly for 3 DWs and then tended to stabilize to the optimal value, *i.e.* around 0.8-mm at 20-mm depth, 1.4-mm at 40-mm depth, 2.1-mm at 60-mm depth and 3.2-mm at 80-mm depth. Moreover, it is worth pointing out that at each depth, the different methods reached the optimal lateral resolution values for the transmission of 15 DWs, leading to an optimal frame rate of 855 fps.

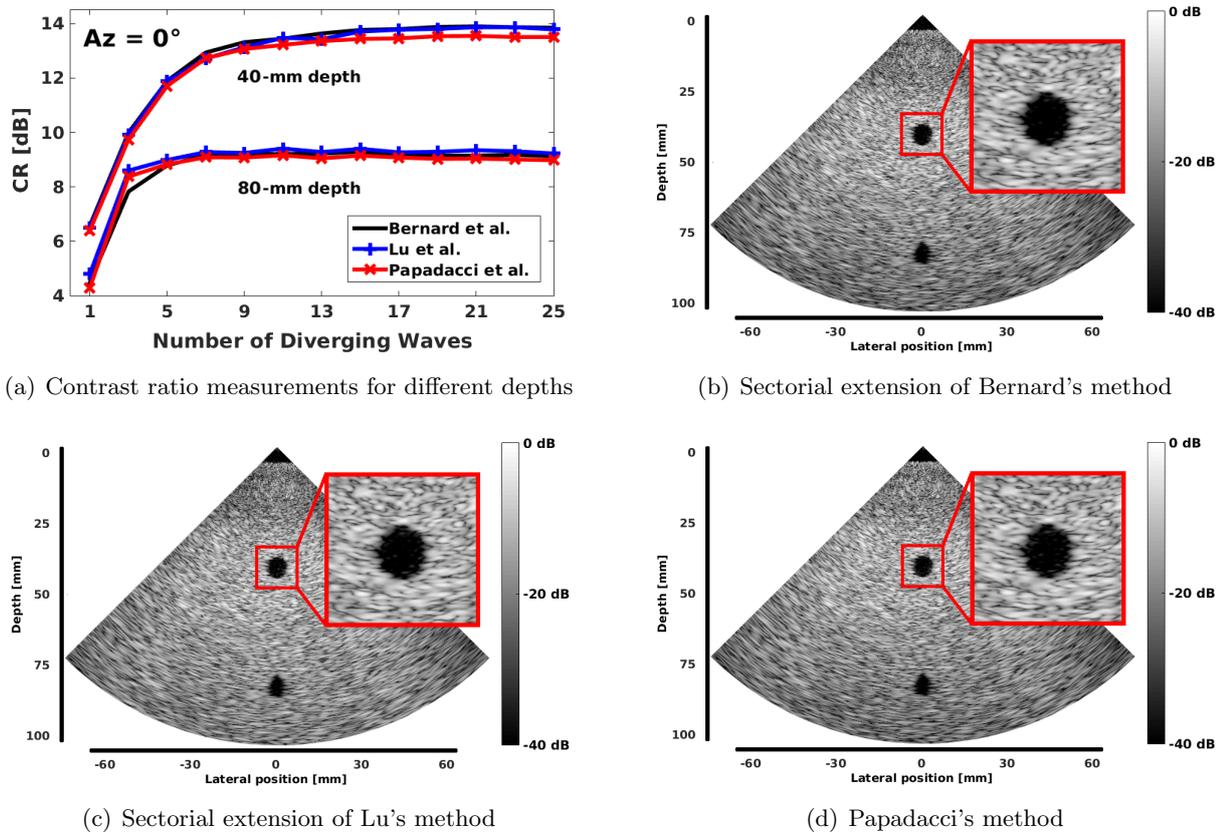


FIGURE 29 – Numerical phantom - (a) contrast ratio measurements computed at the middle of the image (azimuth angle of 0°) as a function of the number of DWs. B-mode images reconstructed from (b) the proposed extension of Bernard's method, (c) the proposed extension of Lu's method and (d) the spatial-based method of Papadacci using 15 DWs with coherent compounding.

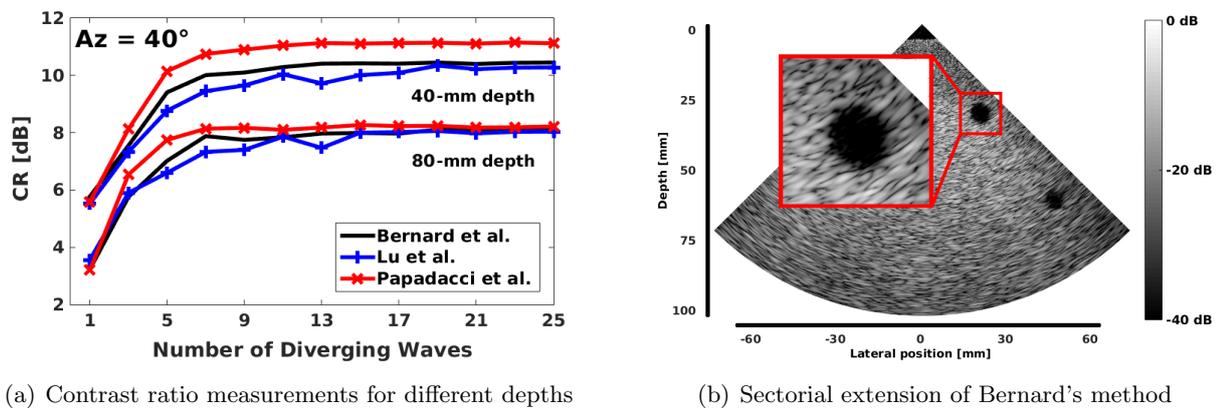


FIGURE 30 – Numerical phantom - (a) contrast ratio measurements computed at the edge of the image (azimuth angle of 40°) as a function of the number of DWs. B-mode images reconstructed from (b) the proposed extension of Bernard's method using 15 DWs with coherent compounding.

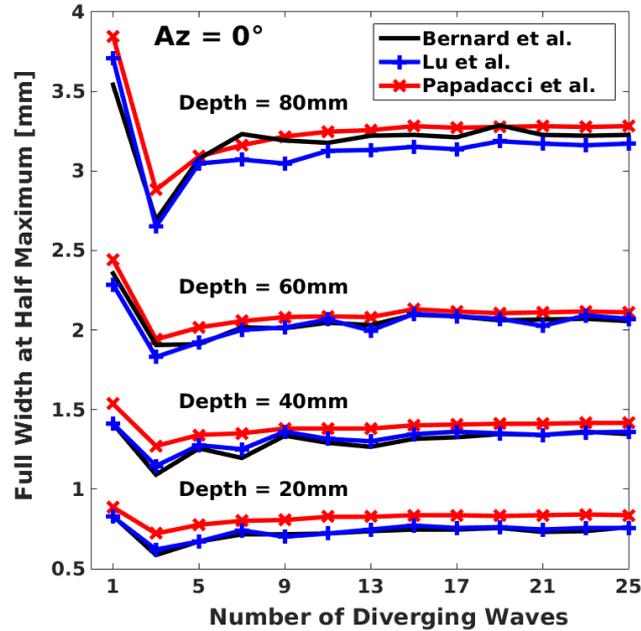


FIGURE 31 – Experimental acquisition - lateral resolution measurements computed at the middle of the image (azimuth angle of 0°) as a function of the number of DWs computed at 20-mm, 40-mm, 60-mm and 80-mm.

Image contrast

We investigated in Fig. 32 the quality of the reconstructed images in terms of contrast. In particular, Fig. 32(a) displays the CR values measured at 40-mm depth for each compounding experiment. From the corresponding graph, it can be observed that the three methods yield very close results. The CR measurements involve same tendency over the increase of the number of DW. The optimal CR values are reached in all cases after the transmission of 15 DW (frame rate of 855 fps), which is in accordance with the simulation results we obtained in Sec. 5.4.1. Fig. 32(b) to 32(d) display the reconstructed B-mode images obtained using 15 DW with coherent compounding. As in simulations, these results illustrate visually the closeness in terms of image quality and speckle definition of the images produced by different methods. It is also interesting to note that the decrease of the intensity observed at the two borders of the image can be explained, on the left side, by the presence of a structure with high reflectivity (due to an unexpected flaw present in the CIRS phantom we used at the interface between the background and the hyperechoic cyst of +3dB) which produces classical ultrasound shadow and, on the right side, by the physical limit of the CIRS phantom which induces reverberation phenomenon. Finally, the limited opening angle of the transducer elements and the limited opening angle of the effective aperture may be additional causes responsible for the observed decrease of intensity at the edges.

In vivo experiment

Fig. 33 displays one particular frame of the full sequence we reconstructed from a real acquisition performed on a healthy volunteer with and without motion compensation, as described in Sec. 5.3.6. The transmission scheme used during the acquisition process allowed us to compute

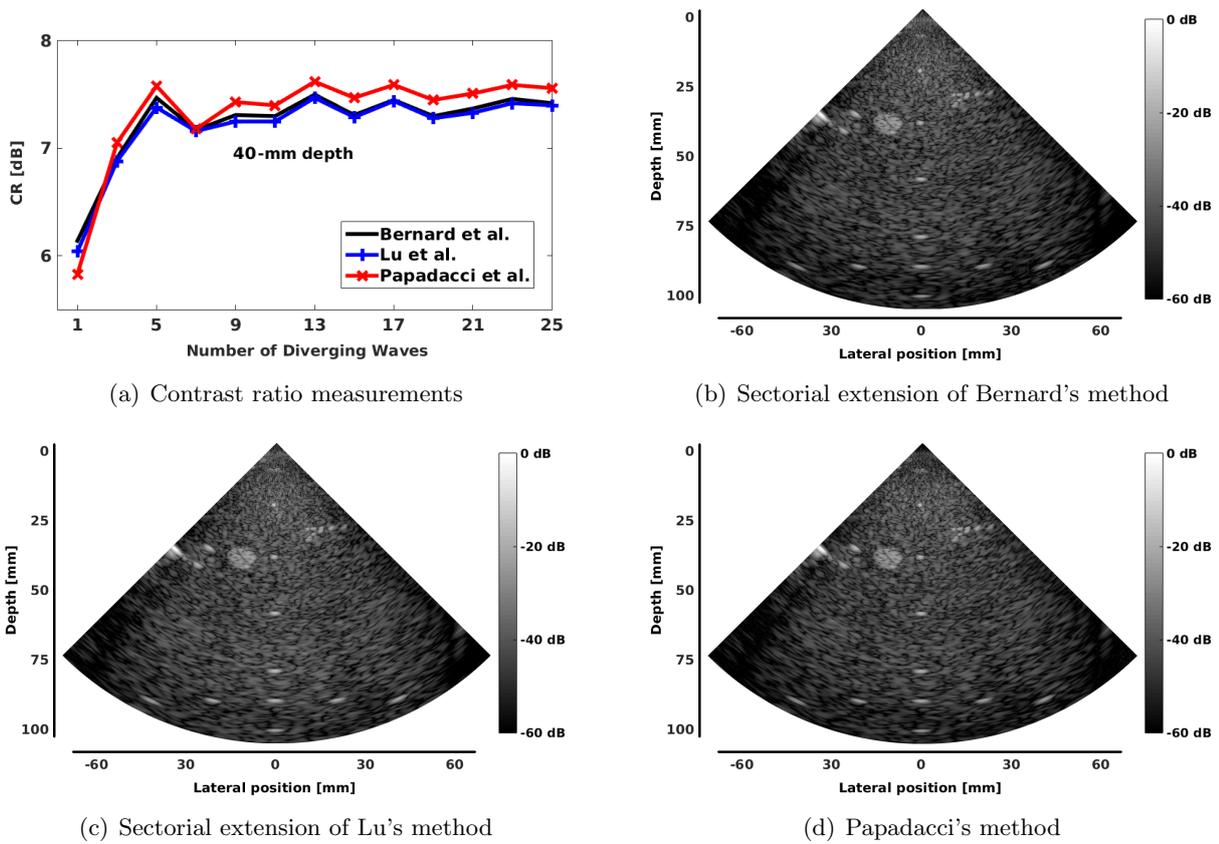


FIGURE 32 – Real acquisition - (a) Contrast ratio measurements as a function of the number of DWs. B-mode images reconstructed from (b) the proposed extension of Bernard's method, (c) the proposed extension of Lu's method and (d) the spatial-based method of Papadacci using 15 DWs with coherent compounding.

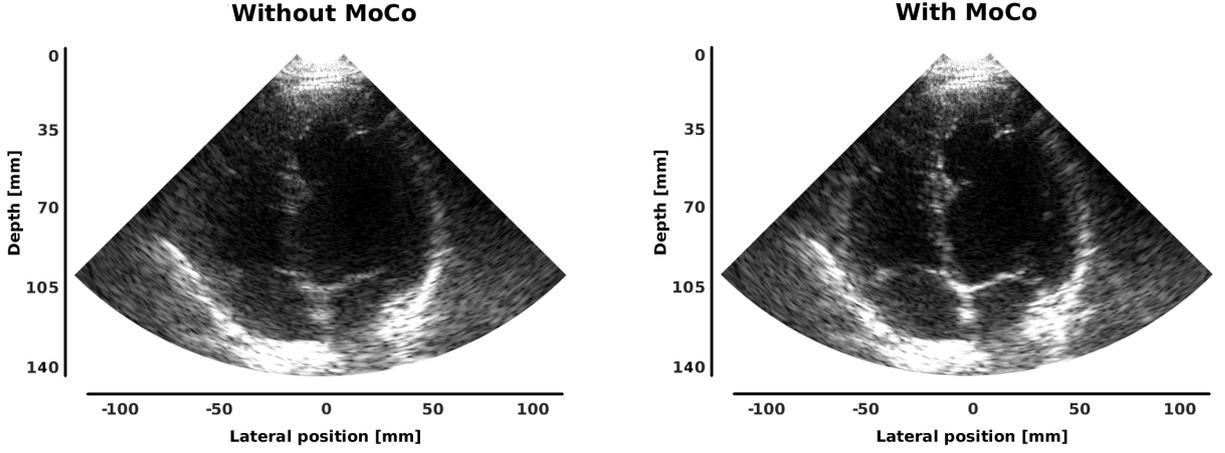


FIGURE 33 – Motion compensation for high frame rate echocardiography of the left ventricle using the sectorial extension of Lu’s method. Each frame has been reconstructed from the transmission of 32 DWs (tilted from -25° to $+25^\circ$ with a triangle strategy as described in [203]). Online movies are also available in the supplementary materials. The data from Porée *et al.* [203] were reanalyzed with the technique described in the present paper.

a B-mode sequence of the whole left and right ventricles at high frame rate (250 fps). This result demonstrates the feasibility of the proposed formulation in a real setting, as well as its flexibility, since we could easily integrate the motion compensation framework proposed by Porée *et al.* [203].

5.5 Discussion

5.5.1 Fourier-based methods of sectorial imaging

We proposed in this study an explicit transformation which allows the extension of existing Fourier-based approaches, initially proposed for planar acquisition, to the reconstruction of sectorial images with wide angle using DW. The proposed formulation was evaluated through the extension of two Fourier-based techniques, *i.e.* the one proposed by Lu *et al.* which samples the Fourier space along the k_z axis direction and the one of Bernard *et al.* which samples the Fourier space radially. Results obtained from both simulations and experiments revealed that the two extended Fourier-based approaches reconstruct images with comparable quality as it is the case for planar acquisition, the introduced spatial transform having no effect on their relative behavior. More importantly, we also demonstrated in this study that the proposed extended Fourier-based approaches produce competitive results compared to the state-of-the-art method proposed by Papadacci *et al.* both in terms of lateral resolution and image contrast. Although the proposed transformation has been specifically designed for the transmission scheme described in [191] (*i.e.* the virtual source point of each DW lies on a horizontal line positioned behind the probe), our approach can also be easily used for other strategies like, for instance, the steered diverging waves presented in [101]. In the case of SDW of angle θ , the second term z_v of Eq. (30) has to be simply replaced by $-\min_E \sqrt{(x_E - x_v)^2 + z_v^2}$ to ensure that the smallest transmit delay is null. The same reasoning then holds to derivate the akin spatial transformation. The corresponding equations have been used to produce the results given in Sec. 5.4.2.

5.5.2 Computational complexity

One strong interest of Fourier-based techniques for sectorial acquisition is a potential reduction of the computational complexity compared to the conventional DAS method. To determine the computational complexity of the proposed extension of Lu and Bernard methods, let n_e denotes the number of elements that compose the transducer (in general 64, 128 or 192), n_s the number of time samples (typically in the range 1000-3000), n_x the number of samples along the x-axis for the reconstructed image (typically between 100 and 500) and n_ξ the number of angles over two used in reception for Bernard's method (the ratio of 2 being explained by the use of half of the spectrum). This value was fixed to 85 in our experiments, the reader is referred to [35] for more details.

Complexity of Papadacci's method

The DAS method proposed by Papadacci *et al.* must retrieve $\lambda(n_x n_s n_e)$ interpolated data and perform $\lambda(n_x n_s)$ summations over n_e , λ being the ratio between the scanned sectorial area and the corresponding encompassing rectangle (with dimensions $n_x n_s$). For an angular aperture of 90° , λ is defined between $[0.5, 1]$. The computational complexity of the method of Papadacci *et al.* thus corresponds to $O(n_x n_s n_e)$.

Complexity of the proposed extension of Lu's method

The sectorial extension of Lu's method requires two 2-D FFTs (complexity of $O(n_x n_s \log(n_x n_s))$) and two times the computation of $(n_x n_s)$ interpolated values (one for the regridding step performed in the Fourier space and one for the spatial transformation introduced in this work). The computational complexity of the extension of Lu's method thus reduces to $O(n_x n_s \log(n_x n_s))$ compared to the method of Papadacci.

Complexity of the proposed extension of Bernard's method

The sectorial extension of Bernard's method requires n_e 1-D FFTs (complexity of $O(n_e n_s \log(n_s))$), one 2D FFT (complexity of $O(n_x n_s \log(n_x n_s))$), the computation of $(n_\xi n_s n_e)$ multiplications for the delays applied in reception and two times the computation of $(n_x n_s)$ interpolated values (one for the regridding step performed in the Fourier space and one for the spatial transformation introduced in this work). The computational complexity of the extension of Bernard's method thus corresponds to $O(n_\xi n_s n_e)$, making the extension of this method less attractive than for the extension of Lu's method.

5.5.3 Extension to temporal acquisitions with tissue motion

Tissue motion is an important source of artifact in ultrafast imaging. In particular, when the motion of the structures of interest between two consecutive firings is higher than $\lambda/8$, it is common to observe a weakening of the compounding effect if no particular strategy is applied. Several approaches have thus been proposed to tackle this problem, both on PW [63] and DW [203]. In this paper, we showed the feasibility of the application of the motion compensation algorithm developed in [203] on the images reconstructed from the proposed approach. This shows the accuracy of our technique and validates further the marginal effect of the error map presented in Sec. 5.2.2.

5.5.4 Perspectives

One potential strong interest of the Fourier-based method compared to the DAS technique concerns its lower computational complexity. In the particular case of sectorial reconstruction, we have shown that the computational ratio between the sectorial extension of Lu's method and the standard DAS technique is of the order $n_e / \log(n_x n_s)$. Although the proposed transformation has been derived for 2D, it can be easily extended to 3D. To this end, one has to take into account the y -coordinate in the travel time equations of the PW (Eq. (32)) and the DW (Eq. (31)). By equating the first order Taylor approximation at $x_E = x_v$ and $y_E = y_v$, one can obtain the corresponding spatial transformation between (x_p, y_p, z_p) and (x_d, y_d, z_d) . Interestingly, in the 3D case, the computational ratio between the proposed method and the DAS technique goes up to $n_{ex} n_{ey} / \log(n_x n_y n_s)$, making the interest of our approach even more stronger.

5.6 Conclusions

In this study, an explicit transformation for the extension of Fourier-based techniques to the reconstruction of sectorial images using DW has been presented. The key concept of the proposed formulation is based on the derivation of an isomorphism in terms of travel time between a planar system based on PW and a sectorial one based on DW. In particular, we proposed an explicit spatial transformation which allows the reconstruction of wide angle images from compounding scheme in transmission, the optimal quality of images being reached for 15 DWs (frame rate of 855 fps) on both numerical and in vitro phantoms. Results obtained from simulations and experiments revealed that the proposed extension of Lu's method produces competitive results with lower computational complexity when compared to conventional delay-and-sum technique.

6 Summary - Conclusions

From Sec. 3 to 5 we presented the work we performed on the reconstruction of ultrasound images. Starting from the observation that Fourier-based approaches are more efficient in terms of computational complexity compared to conventional delay-and-sum techniques, we focused our intention on such approaches. In particular, we developed a novel Fourier-based technique, named Ultrasound Fourier Slice Beamforming, where the acquired samples are radially distributed into the Fourier space. By doing so, the classical Fourier slice theorem can be used to reconstruct ultrasound images. This new technique was validated through both simulation and in vivo acquisitions. Results show that our method produces very competitive results in terms of image quality compared to the state-of-the-art methods but with either equivalent or higher computational complexity, making our solution less attractive. However, from our point of view, the modeling we proposed remains conceptually relevant since this is the first time that strong links between ultrasound image formation with others modalities such as Computed Tomography are created.

Based on the literature review we made in Sec. 2, we developed an innovative method, which is the first Fourier-based approach applicable to diverging wave acquisitions. The key aspects of this approach are the following :

- establishment of an isomorphism in terms of travel time between a planar system based on PW and a sectorial one based on DW ;
- derivation of an explicit spatial transformation which allows the reconstruction of wide-angle sectorial images from standard Fourier-based techniques initially developed for PW imaging ;
- generalization of the explicit spatial transformation from different virtual source positions to allow the use of compounding scheme to improve the quality of the reconstructed image.

Our method was evaluated on both simulations and in vitro/in vivo experiments. Results revealed that the proposed extension of Lu's method produces competitive results with lower computational complexity when compared to conventional delay-and-sum technique. Interestingly, our solution has been successfully applied for the reconstruction of high frame rate cardiac images (250 images/s with a wide sector angle of 90°) from an adult healthy volunteer. This illustrates the ability of our proposed solution to be used for cardiac applications.

My research on ultrasound image reconstruction was carried out through 1 PhD funded by the chinese government (China Scholarship Council grant). Results have led to 4 articles in international journals and 11 articles in international conferences.

Research perspectives

As presented in this manuscript, the estimation of clinical indices in cardiac imaging has improved significantly over the past 15 years. Key publications that we have produced have shown that the most relevant indices used in clinical routines (*i.e.* cavity volumes and ejection fraction) can be automatically estimated with scores below the inter-observer variabilities both in MRI and ultrasound imaging (both in 2D and 3D) [30, 33, 139, 197].

Interestingly, one of the major keys to this improvement is the creation of public referenced databases that integrate more and more patients. Indeed, the design of such databases has first been stimulated by the need for validation linked to the explosion in the number of methods proposed in the field. In particular, thanks to the work we made in [30], we showed that advanced image processing techniques such as our newly developed B-spline Explicit Active Contour approach [197, 213] can provide highly accurate performances (*e.g.* correlations of 0.97, 0.97 and 0.91 for the estimation of the left ventricle volumes at end-diastole, end systole and the ejection fraction in US imaging [24]) with very fast processing time (lower than one second per volume). Interestingly, this study also revealed that machine learning techniques, especially structured random forest [67] and deep learning with anatomical regularization [183], provide among the best current 3-D left ventricle segmentation scores despite the use of a very limited training database (15 patients).

Inspired by the results derived from the CETUS study [30, 183, 197], we investigated in [33, 139] the ability of deep learning solutions to accurately segment cardiac structures based on more advanced databases (patient population of 150 and 500 in MRI and 2D ultrasound, respectively). From both studies, it clearly appears that the problem of cardiac structures segmentation is about to be solved in both modalities, with clinical indices estimated by the best current deep learning networks all below the inter-observer scores and quite close but still higher than those of the intra-observer [33, 139].

As far as cardiac motion estimation is concerned, results derived in our studies are also promising, but clearly less advanced than the ones obtained in cardiac segmentation [7, 299]. The most successful aspects of our work in this area concern the development of a generic pipeline which allows the generation of realistic synthetic cardiac sequences in both MRI (cine/tagged) and ultrasound (2D/3D) imaging [4, 301]. The generated virtual populations have been used to date to evaluate the quality of dedicated approaches (state-of-the-art methods and/or our proposed personalized anatomical deformable model) for estimating cardiac motion and myocardial strain. Results show that the personalized anatomical deformable scheme is a good approach to effectively regularize strain from motion estimation. However, the quality of the derived measures also revealed that there is still a need to improve motion estimation from cardiac imaging. Two serious ways for such improvement are emerging : *i)* the development of dedicated cardiac acquisition sequences to produce new image sequences with better properties

for motion estimation; *ii*) the exploitation of recent advances in deep learning techniques to specialize some of these approaches to motion estimation in cardiac imaging and see whether the same quality of results as those obtained for the segmentation task can be reached. Based on these findings, I plan to organize my research at mid and long term as described below.

Mid-term research perspectives

From a mid-term research point of view, my objective is to continue my work on the reliability of clinical measurements in cardiac imaging.

Cardiac segmentation

With regards to the estimation of robust volumetric indices through cardiac segmentation, my goal is to finish solving this problem by developing dedicated solutions based on deep-learning that will produce results within the intra-observer scores with cardiac shapes which will be anatomically plausible in any case. Indeed, the results we recently obtained in MRI imaging showed that the best performing deep learning method provides on average 1.6 slices out of 12 segmented ones per volume with impossible anatomical shapes (*i.e.* the right ventricle disconnected from the myocardium or the left ventricle cavity in contact with the background) for 41 patients out of the 50 in the testing database [33]. This reveals that even if the most successful current methods obtained on average results within the inter-observer variability, they may still be subject to errors that an expert will not do in clinical routines.

Based on these observations and on the fact that for both MRI and US imaging, the best current methods produce results still higher than the intra-observers scores, I have already begun to explore ways to further improve segmentation results in cardiac imaging. The two serious tracks under consideration concern the localization of structures of interest before segmentation and the introduction of relevant shape constraints into deep learning formalism. Concerning the first aspect, recent work that we have done have shown that accurate localization of the left ventricle allows segmentation scores to be within the intra-observer variability. This stimulated us to work on deep learning architectures that optimize both localization and segmentation in parallel. This work is currently being investigated with one of my PhD student financed by the LABEX PRIMES. With regard to anatomical errors, there is already works in the literature that propose solutions to solve this problem. Among the existing techniques, the work of Oktay *et al.* , recently proposed in [183], seems to be one of the most advanced. In this study, the authors used CNNs to segment the cardiac structures. The core of their neural network is based on an architecture similar to the U-Net, whose segmentation output is constrained to fit a non-linear compact representation of the underlying anatomy derived from an auto-encoder network. The use of an auto-encoder to create a compact dedicated space to efficiently represent cardiac structures seems to be an excellent idea. Unfortunately, the current use of such a space does not guarantee that the segmentation output will be anatomically plausible in any case. I therefore plan to work on such encoder-decoder approaches to generate compact space to efficiently code cardiac shapes but with the hard constraint of imposing plausible anatomical shapes at the output of the segmentation model. Thus, the combination of the two approaches mentioned above will ensure that my final solution will obtain segmentation results within the inter- and intra-observer scores with anatomical shapes that will be relevant in all cases.

Cardiac motion estimation

With regards to the different dynamic indices that can be extracted from cardiac sequences, my objective is first of all to improve the quality of the motion estimation that can be extracted from the images, with a particular focus in ultrasound imaging. Indeed, ultrasound imaging seems to be a modality of choice for motion estimation since it is a real time imaging, with the possibility of making acquisitions at very high rates (> 500 fps) and with an intrinsic speckle texture allowing to follow motion patterns both in muscle and blood.

Based on the very good results we obtained in segmentation, I intend to extend my work on deep learning to motion estimation in cardiac imaging. The current best performing deep learning methods applied in medical imaging are mostly based on supervised learning [147]. Such approaches involve setting up benchmark datasets from which the algorithm learns to reproduce the associated references. The creation of such datasets for applications like view classification or anatomical structures segmentation is “easily” possible. This explains why the majority of deep learning solutions are currently focused on solving such problems, particularly in cardiac imaging [131]. As far as motion estimation is concerned, the establishment of benchmark databases is much more difficult to set up, which explains the reduced number of studies on such subject. Most of the existing deep learning methods which address the motion estimation problem come from computer vision and are based on realistic synthetic sequences to form the training dataset [110]. Thus, the application of such techniques in cardiac imaging will inevitably require, as a first key step, the creation of synthetic datasets that are realistic in terms of motion and image properties. Obviously, I intend to use my work and experience on simulation of realistic synthetic cardiac sequences to generate such benchmark datasets. In particular, I plan to develop a generic deep learning algorithm to estimate any kind of motion in echocardiographic imaging. Such architecture will be fed by a significant number of synthetic data generated through a physical simulator. One of the interest of going through a physical simulator is to integrate the speckle decorrelation phenomenon within the simulated sequences. Indeed this phenomenon is partly responsible for the limited effectiveness of current state-of-the-art methods in estimating motion in echocardiography since it implies differences between the actual tissue motion and the apparent motion on the image. Such a scheme will therefore give the possibility to our deep learning solution to specialize on the particularity of the decorrelation of speckle which happens in ultrasound images and thus move one step forward compared to state-of-the-art methods. It should be noted that the conception of such a physical simulator is already under way through the LABEX PRIMES funding of a postdoctoral project. Indeed, there already exists open access ultrasound simulators among which the most famous is FieldII software. However, this solution is not open-source, coded in matlab and needs a long time to simulate a single image (from several firings). Based on an initiative of my colleague Dr. Garcia, we plan to develop an open access and open-source ultrasound physical simulator which will be efficiently coded in the more appropriated C++ programming language to generate realistic ultrasound sequences in a reasonable time.

In terms of applications, I first plan to develop a deep learning solution dedicated to the estimation of tissue motion in cardiac imaging, with a particular interest in the estimation of myocardial strain. Indeed, many studies have been carried out so far to investigate the possibility of estimating accurate myocardial strain information from echocardiographic imaging [193]. It comes out from these studies that strain quantification has high potential for risk stratification in various cardiac disease but strongly suffers from a lack of reproducibility between vendors, modalities and software version. Based on my recent work on vendor-specific simulation of

realistic synthetic ultrasound data [4], I plan to integrate the image variability between vendors into the design of our synthetic benchmark dataset. An additional key aspect will be the ability of my deep learning solution to provide uncertainty measurements on the estimated motion realized on each pixel. Indeed, a particularity of ultrasound images is that depending on the patients, the acquisition conditions and the experience of cardiologists, the quality of the acquired image varies considerably, which generally leads to a partial lack of information on different regions of the acquired image. For instance, it is traditionally difficult to visualize the anterior wall of the myocardium during an apical two-chamber acquisition due to a loss of signal from the lungs. These missing areas are usually characterized by noise or reconstruction artifacts such as reverberation, clutter or acoustic shadowing. It is therefore essential to be able to locate those regions in order to disregard the corresponding motion estimates when calculating the cardiac strain. I plan to integrate the recent advances in Bayesian deep learning to solve this issue. Indeed, this formalism typically estimates uncertainty by either modeling distributions over weights of the network, or by learning a direct mapping to probabilistic outputs. Recent works have shown the possibility to apply such formalism in medical imaging [234]. However, these are still pilot studies and nothing has been done so far on motion estimation in the field of cardiac imaging. I just started to work on that topic with one PhD student financed by Philips Medisys society through the French CIFRE doctoral scholarships.

A second application I intend to explore is flow estimation in echocardiographic imaging. Indeed, recent studies have investigated the strong link between the myocardium deformation and the underlying flow in the left ventricle [84]. It is obvious that taking these two phenomena into account in an individual and relational way will lead to a better characterization of the type and degree of cardiac pathology of patients and therefore to a better diagnosis. The current limitations of flow estimation in echocardiography is the access to limited size window to perform accurate measurements. Indeed, current clinical applications are based on flow measurements either along straight segments via pulsed Doppler or in small windows via color Doppler (*e.g.* at the mitral ring). Recent advances in high frame rate imaging in ultrasound are opening the door to flow imaging throughout the heart cavity [15]. Nevertheless, the main drawback of these methods concerns the low signal-to-noise ratio which characterizes this type of imaging. The underlying measurements must therefore be strongly regularized to provide meaningful interpretation, to the detriment of the quality and accuracy of the extracted indices. I therefore plan to simulate such low signal-to-noise ratio images obtained from high frame rate reconstruction techniques with appropriate noise sources to design a dedicated deep learning architecture capable of estimating the relevant motion on such ultrasound sequences. In particular, based on a physical simulator, I will generate a synthetic benchmark dataset composed by both b-mode images and color Dopplers. Indeed, I consider that these two kinds of information are complementary and can allow a deep learning algorithm to learn more relevant features for a better estimation of intra-ventricular flow. I just started to work on that topic with one PhD student financed by the LABEX PRIMES.

Long term research perspectives

From a long term research point of view, my objective is to investigate the capacity of deep learning solutions in generating controlled latent space for dedicated medical applications. The interest of building such latent spaces lies in their ability to : *i*) introduce constraints more efficiently into the solution space ; *ii*) model more complex tasks than simply estimating motion/segmentation. Recent advances in deep learning have shown their efficiency in learning

partially-controlled spaces that integrate the complex variability of the studied population. Among the most advanced methods, Variational auto-encoders (VAEs) [128] and adversarial Variational auto-encoders (aVAEs) [156] currently provide the most promising results. Standard auto-encoders encode input data as vectors which allows the derivation of a latent and compressed representation (space) of the data. These methods are useful in dimensionality reduction. Auto-encoders are generally paired with a so-called decoder, which allows to reconstruct input data based on its hidden representation. VAE is a type of auto-encoder with added constraints on latent space. More precisely, instead of letting the auto-encoder learn an arbitrary latent space, VAE imposes a Gaussian distribution a priori centered at zero and of variance one for the latent vector. Thus, a vector sampled from the latent manifold will generate new output data samples, so that VAEs correspond to generative models. One of the main interests of VAEs is therefore their capacity to generate partially-controlled latent spaces defined by the corresponding Gaussian probability distribution. In the aVAEs formalism, the latent space of the VAE is further constrained by an adversarial network that matches the latent vector of the VAE to an arbitrary prior. Thus, by adjusting the a priori probability distribution used by the VAE and the adversarial network involved in the aVAE, it is possible to build partially-controlled latent space which can be tuned for dedicated applications.

Toward deep manifold learning

Based on the previous described formalism, the latent space of the aVAE can be seen as a space approaching a manifold whose properties are directly influenced by the chosen a priori distribution and the adversarial network. It is therefore possible to consider creating manifolds dedicated to tasks that can be too complex to be modelled using conventional approaches. As a proof of concepts, I am currently working in a collaboration with the university of Sherbrooke (Canada) on the exploitation of such a paradigm for an efficient introduction of shape constraints into deep learning framework. However, I am deeply convinced that this formalism is generic enough to be used in much more ambitious medical applications such as : *i*) real-time guidance of the ultrasound probe for better imaging and thus better robustness and reproducibility of the underlying measurements; *ii*) automatic risk assesement of patients with regards to some pathologies based on relevant information extracted from the image or the image itself. Because of their very rich intellectual and application potential, there is no doubt that these prospects will be able to provide my future research with fertile ground in the long term.

French summary

Au cours de mon doctorat, j'ai eu l'occasion de travailler sur la segmentation des images échocardiographiques avec un intérêt particulier pour les statistiques du signal ultrasonore afin d'améliorer l'efficacité du terme d'attache aux données utilisé dans un modèle de contour actif dédié. Ainsi, les principales innovations développées au cours de mon doctorat concernent à la fois une étude complète des statistiques du signal radiofréquence ultrasonore et la dérivation d'une méthode de colocalisation basée sur des fonctions radiales afin de modéliser efficacement l'évolution des ensembles de niveaux lors de la segmentation d'images. Ce travail a été valorisé par la publication de 4 articles (2 comme premier auteur) dans des revues internationales et 6 articles (5 comme premier auteur) dans des conférences internationales.

Les activités de recherche que j'ai menées depuis sont décrites dans les parties ci-dessous. Mon travail concerne le traitement du signal et de l'image appliqué à l'imagerie médicale. En particulier, je me suis concentré sur trois sujets principaux : la segmentation cardiaque, l'estimation du mouvement cardiaque et la reconstruction d'images ultrasonores, le tout pour un meilleur diagnostic des maladies cardiaques. En termes de développements méthodologiques, j'ai travaillé sur *i*) les approches variationnelles au travers du formalisme de surface active explicite B-spline (BEAS) ; *ii*) l'apprentissage statistique au travers des techniques par forêts aléatoires et par apprentissage profond ; *iii*) la modélisation du signal basée sur la théorie du signal monogénique ; *iv*) des problèmes inverses en exploitant le théorème de Fourier ; *v*) des problèmes d'optimisation au travers de régularisation anatomique personnalisée. D'un point de vu applicatif, j'ai consacré mes recherches à l'imagerie cardiaque (principalement l'échographie et l'IRM), à la fois en termes de reconstruction et d'analyse d'images. L'objectif clinique de mes recherches concerne un meilleur diagnostic des maladies cardiaques grâce à des mesures fiables et robustes d'indices cliniques à partir de l'image tels que les volumes ventriculaires, la fraction d'éjection et la déformation myocardique.

J'ai financé mes travaux de recherche via quatre sources principales :

- financements nationaux grâce à des bourses ministérielles (co-encadrement d'un doctorant), au travers du LABEX PRIMES (co-encadrement de deux doctorants et d'un post-doc) et de projets ANR (co-encadrement d'un post-doc via le projet ANR du professeur Hervé Liebgott) ;
- financements européens (co-encadrement de 2 doctorants via la bourse ERC du professeur Jan D'hooge) ;
- financements internationaux (co-encadrement d'un doctorant via une bourse CSC - China Scholarship Council) ;
- financements industriels grâce aux bourses doctorales CIFRE (2 doctorants co-encadrés avec la société Philips Medisys).

1 Résumé de mes travaux de recherche

L'analyse de la fonction cardiaque joue un rôle important en routine clinique pour mesurer la morphologie et la fonction cardiaque et pour établir un diagnostic. Cette analyse est basée sur l'interprétation d'indices cliniques extraits de traitements d'images de bas niveau tels que la segmentation et le suivi. Par exemple, l'extraction de la fraction d'éjection du ventricule gauche nécessite une délimitation précise de la paroi endocardique du ventricule gauche en fin de phase diastolique et en fin de phase systolique. En routine clinique, l'annotation semi-automatique ou manuelle reste un travail quotidien en raison du manque de précision et de reproductibilité des méthodes de segmentation cardiaque entièrement automatisées. Cela conduit à des tâches qui prennent beaucoup de temps et qui sont sujettes à des problèmes de variabilité intra- et inter-experts. Dans ce contexte, mes travaux de recherche se sont concentrés sur la fiabilisation des mesures cliniques en imagerie cardiaque. Je me suis particulièrement intéressé à l'estimation d'indices volumétriques et à l'estimation de la déformation du muscle myocardique. Enfin, j'ai effectué des travaux sur de nouvelles techniques d'imagerie en ultrason afin de créer des systèmes d'imagerie visant une meilleure estimation de mouvement des structures imagées.

Segmentation cardiaque : vers une estimation robuste des indices volumétriques

Afin d'avoir la possibilité d'évaluer le potentiel des méthodes de l'état de l'art en matière de segmentation d'images cardiaques et d'extraction d'indices cliniques pertinents, je me suis tout d'abord consacré à la conception de plusieurs bases de données en accès libre avec des annotations manuelles d'experts (cardiologues et radiologue). J'ai notamment mis en place les bases de données avec les plates-formes d'évaluation en ligne suivantes :

- la base de données CAMUS contenant des séquences échocardiographiques 2D de 500 patients qui ont été acquises lors d'examens cliniques dans le même centre médical avec le même équipement ;
- la base de données CETUS contenant des séquences échocardiographiques 3D de 45 patients qui ont été acquises lors d'examens cliniques dans trois centres différents avec trois équipements différents (GE, Siemens et Philips).
- la base de données ACDC contenant des séquences IRM de 150 patients qui ont été acquises lors d'examens cliniques dans le même centre médical avec deux scanners IRM de forces magnétiques différentes (1.5T et 3.0T). La population cible est composée de patients répartis uniformément en 5 classes avec des caractéristiques bien définies en fonction de paramètres physiologiques.

Grâce à ces bases de données, les innovations et les perspectives suivantes ont été soulignées :

1. CAMUS

- Les réseaux encodeur-décodeur produisent des résultats de segmentation très précis en échocardiographie 2D ;
- Parmi les différentes architectures testées, l'architecture U-Net s'est avérée être la méthode la plus efficace en termes de compromis entre le nombre de paramètres appris et les performances atteintes ;
- L'architecture U-Net a fait preuve d'une robustesse impressionnante vis-à-vis de la variabilité des images ultrasonores. Compte tenu de la grande variété de qualité d'image en imagerie échocardiographique, ce résultat révèle que ce réseau est une solution de choix pour résoudre le problème de segmentation d'images échocardi-

graphiques en 2D ;

- La segmentation et les résultats cliniques des réseaux encodeur-décodeur sont tous compétitifs vis-à-vis des scores inter-observateurs ;
- La segmentation et les résultats cliniques des réseaux encodeur-décodeur sont proches des scores intra-observateurs, mais légèrement moins bons. Même si ces réseaux produisent des résultats remarquables, il y a encore des possibilités d'amélioration afin de reproduire fidèlement l'annotation manuelle d'un expert donné.

2. *CETUS*

- La méthode développée et présentée en section Sec. 4 (BEASM) produit les meilleurs scores actuels pour la segmentation de la paroi endocardique en imagerie échocardiographie 3D ;
- Les plus grandes erreurs produites par la méthode BEASM se produisent au niveau de la région apicale et la paroi antérieure du ventricule gauche ;
- Une méthode d'apprentissage profond avec régularisation anatomique de forme [183] donne des résultats assez proches de ceux obtenus avec notre méthode. L'accès à seulement 15 patients pendant la phase d'entraînement illustre le fort potentiel des techniques d'apprentissage pour l'analyse des images échocardiographiques 3D ;
- les méthodes les plus performantes obtiennent des scores de corrélation précis sur les mesures cliniques avec un biais faible et un écart-type sur les indices LV_{EDV} , LV_{ESV} et LV_{EF} , le LV_{EF} restant les plus difficiles à estimer ;
- La segmentation et les résultats cliniques des méthodes les plus performantes sont tous légèrement inférieurs aux scores inter-observateurs. Ceci révèle qu'il y a encore un besoin d'amélioration de ces méthodes pour la segmentation de la paroi endocardique du ventricule gauche et l'estimation des indices cliniques correspondants en échocardiographie 3D.

3. *ACDC dataset*

- Les réseaux convolutionnels profonds produisent des résultats de segmentation très précis en IRM cardiaque ;
- Les réseaux convolutifs profonds échouent encore à segmenter correctement la base et l'apex, surtout si l'on considère la distance de Hausdorff ;
- Les meilleurs réseaux de neurones convolutionnels obtiennent des scores de corrélation précis sur les mesures cliniques et de faibles biais et écarts-types sur les mesures LV_{EDV} et LV_{EF} , deux des mesures physiologiques les plus utilisées ;
- Les réseaux convolutifs profonds ont encore des difficultés à estimer l'indice RV_{EF} vis-à-vis des autres indices testés ;
- Les résultats cliniques du meilleur réseau neuronal convolutionnel sont tous compétitifs vis-à-vis des scores inter-observateurs ;
- La segmentation et les résultats cliniques du meilleur réseau neuronal convolutionnel sont proches des scores intra-observateurs, mais légèrement moins bon (à l'exception de la mesure de Dice). Ceci révèle que même si les réseaux de neurones convolutifs profonds ont produit des résultats remarquables, il y a encore des possibilités d'amélioration afin de reproduire fidèlement l'annotation manuelle d'un expert donné ;
- Bien que ces observations doivent être validées sur un plus grand nombre de patients, il ressort de cette étude que des techniques d'apprentissage statistiques bien conçues peuvent atteindre des scores de classification presque parfaits.

D'après les résultats obtenus, il apparaît clairement que les méthodes d'apprentissage profond peuvent être considérées comme une solution de choix pour résoudre le problème de segmentation d'images cardiaques et d'estimation d'indices cliniques, tant en échocardiographie 2D et 3D qu'en IRM cardiaque. En effet, lorsque la taille de la base de données est suffisante (c'est-à-dire supérieure à 100 patients dans le cas des études proposées), la segmentation et les résultats cliniques obtenus par ces techniques d'apprentissage sont tous compétitifs vis-à-vis des scores inter-observateurs. Cependant, dans toutes les études réalisées, les réseaux d'apprentissage profond les plus performants n'ont pas réussi à produire des résultats dans la variabilité intra-observateur. Ceci doit stimuler la communauté à améliorer le formalisme d'apprentissage profond en imagerie médicale pour reproduire fidèlement l'annotation manuelle d'un expert sans aucune défaillance anatomique.

Mes travaux de recherche en segmentation cardiaque ont été réalisés au travers de 4 doctorants financés par le gouvernement français (1 bourse ministérielle et 1 bourse LABEX PRIMES) et part l'Europe (2 doctorats financés par la bourse ERC du professeur Jan D'hooge). Les résultats obtenus sur ce thème ont donné lieu à 17 articles dans des revues internationales et 31 articles dans des conférences internationales. J'ai également co-organisé deux challenges sur ce sujet lors de conférences MICCAI (2014 et 2017) et mis en place trois plates-formes en ligne dédiées qui sont toujours ouvertes pour le téléchargement en accès libre des bases de données et à la soumission de nouveaux résultats (<https://www.creatis.insa-lyon.fr/EvaluationPlatform/CETUS/>, <https://acdc.creatis.insa-lyon.fr/> et <http://camus.creatis.insa-lyon.fr/challenge/>)

Estimation du mouvement cardiaque : vers une estimation robuste de la déformation

Partant du constat qu'une validation bien établie de l'estimation de la déformation du muscle myocardique à partir de l'imagerie était nécessaire dans notre communauté, nous avons développé un nouveau pipeline spécialement conçu pour simuler des séquences synthétiques réalistes en imagerie US et IRM à partir d'un même patient virtuel. Le pipeline proposé a été utilisé pour créer un ensemble de données de 18 patients virtuels comprenant des sujets sains et divers cas pathologiques (avec entre autre des cas ischémiques et de dyssynchronie). L'objectif de cet ensemble de données de référence est de fournir à la communauté des outils afin d'établir une comparaison équitable et reproductible des méthodes de l'état de l'art pour l'estimation du mouvement et la quantification de la déformation du muscle myocardique. Les innovations suivantes ont été introduites afin de générer des séquences cardiaques avec un haut degré de réalisme :

- la combinaison d'un modèle électromécanique avec des simulateurs physiques US/IRM afin d'introduire la possibilité d'interagir avec le processus physique de formation d'images ;
- l'utilisation de séquences multimodales réelles d'un même patient afin d'extraire les informations les plus pertinentes pour chaque modalité ;
- l'introduction d'une nouvelle stratégie de déformation combinatoire, dans le but de réduire les artefacts de mouvement qui peuvent se produire dans les régions du myocarde ;
- l'apport du premier cadre unifié qui permet de générer des séquences synthétiques multimodales (US, IRM) réalistes pour un même patient virtuel ;
- l'utilisation d'une nouvelle stratégie basée sur une fonction de pondération gaussienne pour générer une interface lisse entre le myocarde et les structures environnantes.

Nous avons également développé une nouvelle méthode, appelée modèle anatomique déformable (ADM), intégrant les avantages de plusieurs stratégies décrites dans la littérature afin de quantifier efficacement le champ de mouvement et la déformation du muscle myocardique en même temps. La méthode ADM proposée est basée sur les aspects clés suivants :

- l’expression des déformations cardiaques dans le repère anatomique du muscle myocardique. De cette façon, l’intégration de tout mouvement ou régularisation locale sera efficacement adaptée à l’anatomie du coeur ;
- l’utilisation d’une méthode de régression pondérée localement par une approche de partition d’unités afin de régulariser le mouvement et la déformation estimée avec une faible complexité informatique ;
- le renforcement de l’incompressibilité myocardique par la correction des valeurs de déformation dans la direction radiale uniquement.

Grâce à la base de données virtuelles proposée, les innovations et les nouvelles perspectives suivantes ont été soulignées :

- les algorithmes ADM et sparse demons fournissent une meilleure estimation du mouvement et une meilleure quantification de la déformation du myocarde à partir des images IRM marquées et des images US (en comparaison à l’imagerie IRM ciné) ;
- bien que les meilleurs résultats soient obtenus avec l’IRM marquée, les données ultrasonores permettent une estimation précise du mouvement et une quantification de la déformation du myocarde lorsque la qualité de l’acquisition est suffisamment élevée ;
- le modèle ADM semble être la méthode la plus intéressante pour quantifier l’IRM marquée, alors que l’algorithme sparse demons semble mieux convenir à l’imagerie IRM ciné et à l’imagerie ultrasonore ;
- la stratégie de régularisation radiale développée dans l’approche ADM permet une meilleure estimation de la déformation radiale à la fois en imagerie ultrasonore et en imagerie IRM marquée.

Malgré les résultats obtenus encourageants, il apparait que l’estimation du mouvement à partir d’images cardiaques produit des résultats qui ne sont pas encore suffisamment précis et robuste pour être utilisé de façon automatique en routine clinique. En tenant compte des résultats très positifs obtenus récemment en segmentation cardiaque, il semble naturel d’étudier l’efficacité de l’application de solutions d’apprentissage statistique pour l’estimation du mouvement en imagerie cardiaque et de voir si ces approches peuvent être de bons candidats pour résoudre ce problème.

Mes travaux de recherche en estimation du mouvement cardiaque ont été réalisés par le biais d’un doctorant financé par une bourse CIFRE (avec Philips Medisys) et d’un post-doctorant financé par le gouvernement français (financé par le projet ANR du professeur Hervé Liebgott). Les résultats obtenus sur ce thème ont donné lieu à 9 articles dans des revues internationales et 11 articles dans des conférences internationales.

Reconstruction d’images échographiques : vers de nouvelles techniques d’imagerie

Partant du constat que les approches exploitant le formalisme de Fourier sont plus efficaces en termes de complexité informatique que les techniques conventionnelles de délais et sommes, je me suis concentré sur ce type d’approches au cours de mes travaux de recherche. En particulier,

nous avons développé une nouvelle technique, appelée *Ultrasound Fourier Slice Beamforming* - (USTB), où les échantillons acquis sont distribués radialement dans l'espace de Fourier. Ce faisant, le théorème de Fourier classique peut être utilisé afin de reconstruire des images ultrasonores. Cette nouvelle technique a été validée par simulation et par acquisition *in vivo*. Les résultats obtenus montrent que notre méthode produit des résultats très compétitifs en termes de qualité d'image par rapport aux méthodes de pointe, mais avec une complexité de calcul équivalente voire supérieure, ce qui rend notre solution moins attractive. Néanmoins, de mon point de vue, la modélisation que nous avons proposée demeure conceptuellement intéressante puisque c'est la première fois que des liens forts entre la formation d'images échographiques et d'autres modalités comme la tomographie sont établis.

Nous avons également développé une méthode innovante basée sur la théorie de Fourier et applicable aux acquisitions par ondes divergentes. Les principaux aspects de cette approche sont les suivants :

- établissement d'un isomorphisme en termes de temps de parcours entre un système planaire à base d'ondes planes et un système sectoriel à base d'ondes divergentes ;
- dérivation d'une transformation spatiale explicite qui permet la reconstruction d'images sectorielles à grand angle à partir de techniques classiques exploitant la théorie de Fourier et initialement développées pour l'imagerie par ondes planes ;
- généralisation de la transformation spatiale explicite à partir de différentes positions de sources virtuelles pour permettre l'utilisation d'un schéma de composition pour améliorer la qualité des images reconstruites.

Notre méthode a été évaluée à la fois sur des simulations et des expériences *in vitro/in vivo*. Les résultats ont révélé que l'extension proposée de la méthode de Lu produit des résultats compétitifs avec une complexité de calcul plus faible vis-à-vis de la technique conventionnelle de délais et sommes. Fait intéressant, notre solution a été appliquée avec succès pour la reconstruction d'images cardiaques à haute fréquence (250 images/s avec un grand angle d'ouverture de 90°) à partir d'un sujet sain en bonne santé. Ceci illustre la capacité de notre formalisme à être utilisé pour des applications cardiaques.

Mes travaux de recherche en reconstruction d'images échographiques ont été réalisés par le biais d'un doctorant financé par le gouvernement chinois (subvention du China Scholarship Council). Les résultats obtenus sur ce thème ont donné lieu à 4 articles dans des revues internationales et 11 articles dans des conférences internationales.

2 Perspectives de recherche

Perspectives de recherche à moyen terme

Segmentation cardiaque : En ce qui concerne l'estimation d'indices volumétriques robustes par segmentation cardiaque, mon but est de finir de résoudre ce problème en développant des solutions dédiées basées sur les techniques d'apprentissage profond qui produiront des résultats compétitifs vis-à-vis des scores intra-observateurs avec des formes cardiaques anatomiquement plausibles quelque soit les cas de figure. En effet, les résultats que nous avons récemment obtenus en IRM ont montré que la méthode d'apprentissage profond la plus performante fournit en moyenne 1,6 coupes (sur 12 segmentées par volume) avec des formes anatomiques fausses pour 41 patients sur 50 de la base de données des tests. Concernant la précision de segmentation, des travaux récents que nous avons réalisés ont montré que la localisation précise du ventricule

gauche permet d'obtenir des scores de segmentation compétitifs vis-à-vis de la variabilité intra-observateur. Ceci nous incite à travailler sur des architectures d'apprentissage profond qui optimisent à la fois la localisation et la segmentation en parallèle. Ce travail est actuellement à l'étude avec un de mes doctorants financé par les LABEX PRIMES. En ce qui concerne les erreurs anatomiques, il existe déjà dans la littérature des travaux qui proposent des solutions pour résoudre ce problème. Parmi les techniques existantes, le travail d'Oktay *et al.*, récemment proposé dans [183], semble être le plus avancé. Dans cette étude, les auteurs ont utilisé des réseaux convolutionnels pour segmenter les structures cardiaques. Le coeur de leur réseau est basé sur une architecture similaire à celle de U-Net, mais dont la sortie de segmentation est contrainte par une représentation compacte non linéaire de l'anatomie sous-jacente dérivée d'un réseau auto-encodeur. L'utilisation d'un auto-encodeur pour créer un espace dédié compact afin de représenter efficacement les structures cardiaques semble être une excellente idée. Malheureusement, l'utilisation actuelle d'un tel espace ne garantit pas que la sortie de segmentation sera anatomiquement plausible dans tous les cas. J'envisage donc de travailler sur de telles approches encodeur-décodeur pour générer un espace compact afin de coder efficacement les formes cardiaques mais avec la contrainte d'imposer des formes anatomiques plausibles à la sortie du modèle de segmentation. Ainsi, la combinaison des deux approches mentionnées ci-dessus assurera que ma solution finale obtiendra des résultats de segmentation compétitifs vis-à-vis des scores inter- et intra-observateurs avec des formes anatomiques qui seront pertinentes dans tous les cas.

Estimation du mouvement cardiaque : En ce qui concerne les différents indices dynamiques qui peuvent être extraits des séquences cardiaques, mon objectif est tout d'abord d'améliorer la qualité de l'estimation du mouvement qui peut être extrait des images, avec une attention particulière en imagerie ultrasonore. En effet, l'imagerie par ultrasons semble être une modalité de choix pour l'estimation du mouvement puisqu'il s'agit d'une imagerie en temps réel, avec la possibilité de faire des acquisitions à des fréquences très élevées (> 500 fps) et avec des propriétés de speckle intrinsèque à cette imagerie permettant de suivre des modèles de mouvement dans le muscle et le sang. Sur la base des très bons résultats que nous avons obtenus en segmentation, j'ai l'intention d'étendre mes travaux en apprentissage profond à l'estimation du mouvement en imagerie cardiaque. Les méthodes d'apprentissage profond les plus performantes actuellement en imagerie médicale sont principalement basées sur l'apprentissage supervisé [147]. De telles approches impliquent la mise en place d'ensembles de données de référence à partir desquels l'algorithme apprend à reproduire les références associées. La création de tels ensembles de données pour des applications telles que la classification des vues ou la segmentation des structures anatomiques est "facilement" réalisable. C'est pourquoi la majorité des solutions en apprentissage profond se concentrent actuellement sur la résolution de ces problèmes, en particulier en imagerie cardiaque [131]. En ce qui concerne l'estimation du mouvement, l'établissement de bases de données de référence est beaucoup plus difficile à mettre en place, ce qui explique le nombre réduit d'études sur ce sujet. La plupart des méthodes d'apprentissage profond existantes qui abordent le problème de l'estimation du mouvement proviennent de la vision par ordinateur et sont basées sur des séquences synthétiques réalistes pour former la base d'apprentissage [110]. Ainsi, l'application de ces techniques en imagerie cardiaque nécessitera inévitablement, comme première étape clé, la création d'ensembles de données synthétiques réalistes en termes de mouvement et de propriétés d'image. Évidemment, j'ai l'intention d'utiliser mes travaux en simulation pour générer de tels ensembles de données de référence. En particulier, j'envisage de développer un algorithme générique d'apprentissage profond pour estimer tout type de mouvement en imagerie échocardiographique. Cette architecture sera alimentée par un nombre important de données synthétiques générées par un simulateur physique. L'un des intérêts

d'utiliser des simulateurs physiques est d'intégrer le phénomène de décorrélation du speckle dans les séquences simulées. En effet, ce phénomène est en partie responsable de l'efficacité limitée des méthodes actuelles dans l'estimation du mouvement en échocardiographie puisqu'il implique des différences entre le mouvement réel du tissu et le mouvement apparent sur l'image. Un tel schéma donnera donc la possibilité à notre solution d'apprentissage profond de se spécialiser sur la particularité de la décorrélation du speckle et ainsi faire un pas en avant vis-à-vis des méthodes les plus modernes. Il est à noter que la conception d'un tel simulateur physique est déjà en cours grâce au financement par le LABEX PRIMES d'un projet postdoctoral. En effet, il existe déjà des simulateurs d'échographie à accès libre dont le plus célèbre est le logiciel FieldII. Cependant, cette solution n'est pas open-source, codée en matlab et nécessite beaucoup de temps pour simuler une seule image (à partir de plusieurs tirs en émission). A l'initiative de mon collègue Dr Garcia, nous sommes entrain de développer un simulateur physique d'ultrasons, codé efficacement dans le langage de programmation C++ approprié pour générer des séquences d'ultrasons réalistes dans un temps raisonnable. De plus, le code source de ce simulateur sera libre.

En termes d'applications, j'envisage tout d'abord de développer une solution d'apprentissage profond dédiée à l'estimation du mouvement des tissus en imagerie cardiaque, avec un intérêt particulier pour l'estimation de la déformation myocardique. En effet, de nombreuses travaux ont été menées jusqu'à présent afin d'étudier la possibilité d'estimer des informations précises sur la déformation myocardique à partir de l'imagerie échocardiographique [193]. Il ressort de ces études que la quantification de la déformation myocardique présente un potentiel élevé d'évaluation du risque dans diverses maladies cardiaques, mais souffre fortement d'un manque de reproductibilité entre les constructeurs, les modalités et la version du logiciel. En me basant sur mes travaux récents sur la simulation de données échographiques synthétique spécifiques aux constructeurs [4], j'envisage d'intégrer la variabilité des images entre les constructeurs dans la conception de notre ensemble de données de référence synthétique. Un autre aspect clé sera la capacité de ma solution d'apprentissage profond à fournir des mesures d'incertitude sur le mouvement estimé réalisé sur chaque pixel. En effet, une particularité des images échographiques est que selon les patients, les conditions d'acquisition et l'expérience des cardiologues, la qualité de l'image acquise varie considérablement, ce qui conduit généralement à un manque partiel d'information sur différentes régions de l'image acquise. Par exemple, il est traditionnellement difficile de visualiser la paroi antérieure du myocarde lors d'une acquisition apicale deux chambres en raison d'une perte de signal due aux poumons. Ces zones manquantes sont habituellement caractérisées par du bruit ou des artefacts de reconstruction comme la réverbération, ou l'ombre acoustique. Il est donc essentiel de pouvoir localiser ces régions afin de ne pas tenir compte des estimations de mouvement correspondantes lors du calcul de la déformation cardiaque. J'ai l'intention d'intégrer les progrès récents de l'apprentissage profond bayésien pour résoudre ce problème. En effet, ce formalisme permet d'estimer une incertitude soit en modélisant des distributions sur les poids du réseau, soit en apprenant une correspondance directe avec les résultats probabilistes. Des travaux récents ont montré la possibilité d'appliquer ce formalisme à l'imagerie médicale [234]. Cependant, il s'agit encore d'études pilotes et rien n'a été fait jusqu'à présent sur l'estimation du mouvement dans le domaine de l'imagerie cardiaque. Je viens de débiter des travaux sur ce sujet avec un doctorant financé par la société Philips Medisys au travers d'une bourse doctorale CIFRE.

Une deuxième application que j'ai l'intention d'explorer est l'estimation du flux en imagerie échocardiographique. En effet, des travaux récents ont étudié le lien étroit entre la déformation

du myocarde et l'écoulement sous-jacent dans le ventricule gauche [84]. Il est évident que la prise en compte de ces deux phénomènes de manière individuelle et relationnelle conduira à une meilleure caractérisation du type et du degré de pathologie cardiaque des patients et donc à un meilleur diagnostic. Les limites actuelles de l'estimation du flux en échocardiographie sont l'accès à une fenêtre de taille limitée pour effectuer des mesures précises. En effet, les applications cliniques actuelles sont basées sur des mesures de flux soit le long de direction de tirs par Doppler pulsé, soit dans de petites fenêtres par Doppler couleur (*e.g.* au niveau de l'anneau mitral). Les progrès récents de l'imagerie à haute fréquence d'images en échographie ouvrent la porte à l'imagerie de flux dans toute la cavité cardiaque [15]. Néanmoins, le principal inconvénient de ces méthodes est le faible rapport signal/bruit qui caractérise ce type d'imagerie. Les mesures sous-jacentes doivent donc être fortement régularisées pour fournir une interprétation significative, au détriment de la qualité et de la précision des indices extraits. J'ai donc l'intention de simuler des images à faible rapport signal/bruit obtenues à partir de techniques de reconstruction à haute fréquence d'images avec des sources de bruit appropriées pour concevoir une architecture d'apprentissage profond dédiée capable d'estimer le mouvement pertinent sur ces séquences ultrasonores. En particulier, à partir d'un simulateur physique, je vais générer un ensemble de données synthétique de référence composé à la fois d'images en mode B et de Doppler couleur. En effet, je considère que ces deux types d'informations sont complémentaires et peuvent permettre à un algorithme d'apprentissage profond d'apprendre des caractéristiques plus pertinentes pour une meilleure estimation du flux intra-ventriculaire. Je viens de commencer à travailler sur ce sujet avec une doctorante financée par le LABEX PRIMES.

Perspectives de recherche à long terme

D'un point de vue de recherche à long terme, mon objectif est d'étudier la capacité des solutions d'apprentissage profond à générer un espace latent contrôlé pour des applications médicales dédiées. L'intérêt de construire de tels espaces latents réside dans leur capacité à : *i*) introduire plus efficacement des contraintes dans l'espace de solution ; *ii*) modéliser des tâches plus complexes que la simple estimation du mouvement/segmentation. Les progrès récents en apprentissage profond ont démontré leur efficacité dans l'apprentissage d'espaces partiellement contrôlés qui intègrent la variabilité complexe de la population étudiée. Parmi les méthodes les plus avancées, les auto-encodeurs variationnels (VAE) [128] et les auto-encodeurs contradictoires (aVAEs) [156] fournissent actuellement les résultats les plus prometteurs. Les auto-encodeurs standard codent les données d'entrée sous forme de vecteurs, ce qui permet de dériver une représentation latente et comprimée (espace) des données. Ces méthodes sont utiles pour réduire la dimensionnalité. Les auto-encodeurs sont généralement couplés à un décodeur, ce qui permet de reconstruire les données d'entrée en fonction de leur représentation cachée. La méthode dite VAE est un type d'auto-encodeurs avec des contraintes supplémentaires sur l'espace latent. Plus précisément, au lieu de laisser l'auto-encodeurs apprendre un espace latent arbitraire, les VAEs imposent une distribution gaussienne a priori centrée à zéro et de variance un pour le vecteur latent. Ainsi, un vecteur échantillonné à partir du espace latent générera de nouveaux échantillons de données de sortie, de sorte que les VAEs correspondent à un modèle générateur. L'un des principaux intérêts des VAEs est donc leur capacité à générer des espaces latents partiellement contrôlés définis par la distribution de probabilité gaussienne correspondante. Dans le formalisme des aVAEs, l'espace latent des VAEs est contraint par un réseau contradictoire qui fait coïncider le vecteur latent correspondant à un priori arbitraire. Ainsi, en ajustant la distribution de probabilité a priori utilisée par la VAE et le réseau contradictoire utilisé dans les aVAEs, il est possible de construire un espace latent partiellement contrôlé qui peut être adapté

pour des applications spécifiques.

Vers un apprentissage profond de manifolds dédiés : Sur la base du formalisme décrit précédemment, l'espace latent des aVAEs peut être considéré comme un espace approchant un manifold dont les propriétés sont directement influencées par la distribution a priori choisie et le réseau contradictoire. Il est donc possible d'envisager la création de manifolds dédiés à des tâches qui peuvent être trop complexes pour être modélisées par des approches conventionnelles. Comme preuve de concepts, je travaille actuellement en collaboration avec l'université de Sherbrooke (Canada) sur l'exploitation d'un tel paradigme pour une introduction efficace des contraintes de forme dans un cadre d'apprentissage profond. Cependant, je suis profondément convaincu que ce formalisme est suffisamment générique pour être utilisé dans des applications médicales beaucoup plus ambitieuses telles que : *i*) le guidage en temps réel d'une sonde échographique pour une meilleure imagerie et donc une meilleure robustesse et reproductibilité des mesures sous-jacentes ; *ii*) l'évaluation automatique du risque des patients par rapport à certaines pathologies basées sur l'image même. En raison de leur potentiel intellectuel et applicatif très riche, il n'y a pas de doute que ces perspectives pourront fournir à mes travaux de recherche futurs un terrain fertile sur du long terme.

Bibliography

- [1] *Computerized Imaging Reference Systems Inc.* [Online]. Available : <http://www.cirsinc.com/products/all/80/general-purpose-ultrasound-phantom/>, 2013.
- [2] *The CGAL Project*, CGAL User and Reference Manual. CGAL Editorial Boards, [Online]. Available : <http://doc.cgal.org/4.9/Manual/packages.html>, 2016.
- [3] M. ALESSANDRINI, A. BASARAB, H. LIEBGOTT et O. BERNARD : Myocardial motion estimation from medical images using the monogenic signal. *IEEE Transactions on Image Processing*, 22(3):1084–1095, 2013.
- [4] M. ALESSANDRINI, B. CHAKRABORTY, B. HEYDE, O. BERNARD, M. De CRAENE, M. SERMESANT et J. D’HOOGHE : Realistic vendor-specific synthetic ultrasound data for quality assurance of 2-d speckle tracking echocardiography : Simulation pipeline and open access database. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 65(3):411–422, March 2018. ISSN 0885-3010.
- [5] M. ALESSANDRINI, M. De CRAENE, O. BERNARD, S. GIFFARD-ROISIN, P. ALLAIN, I. WAECHTER-STEHLER, J. WEESE, E. SALOUX, H. DELINGETTE, M. SERMESANT *et al.* : A pipeline for the generation of realistic 3d synthetic echocardiographic sequences : methodology and open-access database. *IEEE transactions on medical imaging*, 34(7):1436–1451, 2015.
- [6] M. ALESSANDRINI, D. FRIBOULET, O. BASSET, J. D’HOOGHE et O. BERNARD : Level-set segmentation of myocardium and epicardium in ultrasound images using localized bhattacharyya distance. *In 2009 IEEE International Ultrasonics Symposium*, pages 2468–2471, Sept 2009.
- [7] M. ALESSANDRINI, B. HEYDE, S. QUEIRÓS, S. CYGAN, M. ZONTAK, O. SOMPHONE, O. BERNARD *et al.* : Detailed evaluation of five 3d speckle tracking algorithms using synthetic echocardiographic recordings. *IEEE Transactions on Medical Imaging*, 35(8):1915–1926, 2016.
- [8] M. ALESSANDRINI, H. LIEBGOTT, D. FRIBOULET et O. BERNARD : Simulation of realistic echocardiographic sequences for ground-truth validation of motion estimation. *In 2012 19th IEEE International Conference on Image Processing*, pages 2329–2332, Sept 2012.
- [9] N. ALMEIDA, D. FRIBOULET, S. I. SARVARI, O. BERNARD, D. BARBOSA, E. SAMSET et J. D’HOOGHE : Left-Atrial Segmentation From 3-D Ultrasound Using B-Spline Explicit Active Surfaces With Scale Uncoupling. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 63(2):212–221, Feb 2016. ISSN 0885-3010.

- [10] A. A. AMINI, Yasheng CHEN, M. ELAYYADI et P. RADEVA : Tag surface reconstruction and tracking of myocardial beads from spamm-mri with parametric b-spline surfaces. *IEEE Transactions on Medical Imaging*, 20(2):94–103, 2001.
- [11] E. ANGELINI, S. HOMMA, G. PEARSON, J. HOLMES et A. LAINE : Segmentation of real-time three-dimensional ultrasound for quantification of ventricular function : A clinical study on right and left ventricles. *Ultrasound In Medicine and Biology*, 31(9):1143–1158, 2005.
- [12] G. D. AQUARO, G. CAMASTRA, L. MONTI, M. LOMBARDI, A. PEPE, S. CASTELLETTI *et al.* : Reference values of cardiac volumes, dimensions, and new functional parameters by mr : A multicenter, multivendor study. *Journal of Magnetic Resonance Imaging*, 45(4):1055–1067, 2016.
- [13] C. ARMSTRONG, E. P. RICKETTS, C. COX, P. ADLER, A. ARYNCHYN *et al.* : Quality Control and Reproducibility in M-Mode, Two-Dimensional, and Speckle Tracking Echocardiography Acquisition and Analysis : The CARDIA Study, Year 25 Examination Experience. *Echocardiography*, 32(8):1233–1240, 2015.
- [14] T. ARTS, F. PRINZEN, T. DELHAAS, J. MILLES, A. ROSSI et P. CLARYSSE : Mapping displacement and deformation of the heart with local sine-wave modeling. *IEEE transactions on medical imaging*, 29:1114–23, 03 2010.
- [15] Kondo Claude ASSI, Etienne GAY, Christophe CHNAFA, Simon MENDEZ, Franck NICLOUD, Juan F P J ABASCAL, Pierre LANTELME, François TOURNOUX et Damien GARCIA : Intraventricular vector flow mapping—a doppler-based regularized problem with automatic model selection. *Physics in Medicine & Biology*, 62(17):7131–7147, aug 2017.
- [16] M. R. AVENDI, A. KHERADVAR et H. JAFARKHANI : A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI. *Medical Image Analysis*, 30:108–119, 2016.
- [17] L AXEL et L DOUGHERTY : Mr imaging of motion with spatial modulation of magnetization. *Radiology*, 171(3):841–845, 1989.
- [18] I. Ben AYED, H. M. CHEN, K. PUNITHAKUMAR, I. ROSS et S. LI : Max-flow segmentation of the left ventricle by recovering subject-specific distributions via a bound of the bhattacharyya measure. *Medical Image Analysis*, 16(1):87–100, 2012.
- [19] W. BAI, W. SHI, C. LEDIG et D. RUECKERT : Multi-atlas segmentation with augmented features for cardiac MR images. *Medical Image Analysis*, 19(1):98–109, 2015.
- [20] P. BARALDI, A. SARTI, C. LAMBERTI, A. PRANDINI et F. SGALLARI : Evaluation of differential optical flow techniques on synthesized echo images. *IEEE Transactions on Biomedical Engineering*, 43(3):259–272, March 1996. ISSN 0018-9294.
- [21] D. BARBOSA, O. BERNARD, B. HEYDE, T. DIETENBECK, D. FRIBOULET et J. D’HOOGHE : Hybrid Energy Approach for Real-time B-spline Explicit Active Tracking of Surfaces (heartBEATS). *In Proc. of IUS2013*, pages 663 – 666, 2013.

-
- [22] D. BARBOSA, T. DIETENBECK, B. HEYDE, H. HOULE, D. FRIBOULET, J. D’HOOGHE et O. BERNARD : Fast and Fully Automatic 3-D Echocardiographic Segmentation Using B-Spline Explicit Active Surfaces : Feasibility Study and Validation in a Clinical Setting. *Ultrasound in Medicine and Biology*, 39(1):89 – 101, 2013.
- [23] D. BARBOSA, T. DIETENBECK, B. HEYDE, H. HOULE, D. FRIBOULET, J. D’HOOGHE et O. BERNARD : Fast and fully automatic 3D echocardiographic segmentation using b-spline explicit active surfaces : Feasibility study and validation in a clinical setting. *Ultrasound In Medicine and Biology*, 39(1):89–101, 2013.
- [24] D. BARBOSA, T. DIETENBECK, B. HEYDE, H. HOULE, D. FRIBOULET, J. D’HOOGHE et **O. Bernard** : Fast and Fully Automatic 3-D Echocardiographic Segmentation Using B-Spline Explicit Active Surfaces : Feasibility Study and Validation in a Clinical Setting. *Ultrasound in Medicine and Biology*, 39(1):89 – 101, 2013. ISSN 0301-5629. URL <http://www.sciencedirect.com/science/article/pii/S0301562912004723>.
- [25] D. BARBOSA, T. DIETENBECK, J. SCHAEERER, J. D’HOOGHE, D. FRIBOULET et O. BERNARD : B-Spline Explicit Active Surfaces : An Efficient Framework for Real-Time 3-D Region-Based Segmentation. *IEEE Transactions on Image Processing*, 21(1):241–251, 2012.
- [26] D. BARBOSA, D. FRIBOULET, J. D’HOOGHE et O. BERNARD : Fast tracking of the left ventricle using global anatomical affine optical flow and local recursive block matching. *In Proc. MICCAI Challenge on Echocardiographic Three-Dimensional Ultrasound Segmentation (CETUS), Boston, MIDAS Journal*, pages 17–24. 2014.
- [27] D. BARBOSA, D. FRIBOULET, J. D’HOOGHE et O. BERNARD : Fast tracking of the left ventricle using global anatomical affine optical flow and local recursive block matching. *In Proceedings MICCAI Challenge on Echocardiographic Three-Dimensional Ultrasound Segmentation (CETUS), Boston, MIDAS Journal*, pages 17–24, 2014.
- [28] C. BAUMGARTNER, L. M KOCH, M. POLLEFEYS et E. KONUKOGLU : An exploration of 2d and 3d deep learning techniques for cardiac mr image segmentation. *In Proc. STACOM-MICCAI, LNCS, volume 10663*, pages 111–119, 2017.
- [29] J. BERCOFF, G. MONTALDO, T. LOUPAS, D. SAVERY, F. MÉZIÈRE, M. FINK et M. TANTER : Ultrafast compound doppler imaging : providing full blood flow characterization. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 58(1):134–147, 2011.
- [30] O. BERNARD, J. G. BOSCH, B. HEYDE, M. ALESSANDRINI, D. BARBOSA, S. CAMARASU-POP, F. CERVENANSKY, S. VALETTE, O. MIREA *et al.* : Standardized Evaluation System for Left Ventricular Segmentation Algorithms in 3D Echocardiography. *IEEE Transactions on Medical Imaging*, 35(4):967–977, 2016.
- [31] O. BERNARD, J. D’HOOGHE et D. FRIBOULET : Statistics of the radio-frequency signal based on K distribution with application to echocardiography. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 53(9):1689–1694, Sept 2006. ISSN 0885-3010.

- [32] O. BERNARD, B. HEYDE, M. ALESSANDRINI, D. BARBOSA, S. CAMARASU-POP, F. CERVENANSKY, S. VALETTE, O. MIREA, E. GALLI, M. GELEIJNSE, A. PAPACHRISTIDIS, J.G. BOSCH et J. D’HOOGHE : Challenge on endocardial three-dimensional ultrasound segmentation (cetus). *In Proceedings MICCAI Challenge on Echocardiographic Three-Dimensional Ultrasound Segmentation (CETUS), Boston, MIDAS Journal*, pages 1–8, 2014.
- [33] O. BERNARD, A. LALANDE, C. ZOTTI, F. CERVENANSKY, X. YANG, P. HENG, I. CETIN, K. LEKADIR, O. CAMARA, M. A. G. BALLESTER, G. SANROMA, S. NAPEL, S. PETERSEN, G. TZIRITAS, E. GRINIAS, M. KHENED, V. A. KOLLERATHU, G. KRISHNAMURTHI, M. ROHÉ, X. PENNEC, M. SERMESANT, F. ISENSEE, P. JAGER, K. H. MAIER-HEIN, C. F. BAUMGARTNER, L. M. KOCH, J. M. WOLTERINK, I. ISGUM, Y. JANG, Y. HONG, J. PATRAVALI, S. JAIN, O. HUMBERT et P. M. JODOIN : Deep Learning Techniques for Automatic MRI Cardiac Multi-structures Segmentation and Diagnosis : Is the Problem Solved? *IEEE Transactions on Medical Imaging*, 37(11):2514–2525, 2018.
- [34] O. BERNARD, B. TOUIL, J. D’HOOGHE et D. FRIBOULET : Statistical Modeling of the Radio-Frequency Signal for Partially- and Fully-Developed Speckle Based on a Generalized Gaussian Model with Application to Echocardiography. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 54(10):2189–2194, October 2007. ISSN 0885-3010.
- [35] O. BERNARD, M. ZHANG, F. VARRAY, P. GUETH, J.P. THIRAN, H. LIEBGOTT et D. FRIBOULET : Ultrasound fourier slice imaging : a novel approach for ultrafast imaging technique. *In IEEE International Ultrasonics Symposium (IUS)*, pages 129–132. Ieee, 2014.
- [36] M. BERNIER, P. M. JODOIN et A. LALANDE : Automatized evaluation of the left ventricular ejection fraction from echocardiographic images using graph cut. *In Proc. MICCAI Challenge on Echocardiographic Three-Dimensional Ultrasound Segmentation (CETUS), Boston, MIDAS Journal*, pages 25–32. 2014.
- [37] P. J. BESL et N. D. MCKAY : A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, Feb 1992. ISSN 0162-8828.
- [38] H. BLESSBERGER et T. BINDER : Two dimensional speckle tracking echocardiography : basic principles. *Heart*, 96(9):716–722, 2010. ISSN 1355-6037. URL <https://heart.bmj.com/content/96/9/716>.
- [39] J. BOGAERT, S. DYMARKOWSKI, A.M. TAYLOR et V. MUTHURANGU : Cardiac Function. *In Clinical Cardiac MRI*, pages 109–168. Springer, 2012.
- [40] S. BOGARAPU, M. D. PUCHALSKI, M. EVERITT, R. V. WILLIAMS, H. Y. WENG et S. MENON : Novel cardiac magnetic resonance feature tracking (cmr-ft) analysis for detection of myocardial fibrosis in pediatric hypertrophic cardiomyopathy. *Pediatric Cardiology*, 37, 01 2016.
- [41] G. CARNEIRO, J. C. NASCIMENTO et A. FREITAS : The Segmentation of the Left Ventricle of the Heart From Ultrasound Data Using Deep Learning Architectures and Derivative-Based Search Methods. *IEEE Transactions on Image Processing*, 21(3):968–982, 2012.

-
- [42] M. D. CERQUEIRA, N. J. WEISSMAN, V. DILSIZIAN, A. K. JACOBS, S. KAUL, W. K. LASKEY, D. J. PENNELL, J. A. RUMBERGER, T. RYAN et M. S. VERANI : Standardized Myocardial Segmentation and Nomenclature for Tomographic Imaging of the Heart. *Circulation*, 105(4):539–542, 2002. ISSN 0009-7322.
- [43] I. CETIN, G. SANROMA, S. E. PETERSEN, S. NAPEL, O. CAMARA, M. A. González BALLESTER et K. LEKADIR : A radiomics approach to computer-aided diagnosis in cardiac cine-mri. *In Proc. STACOM-MICCAI, LNCS, volume 10663*, pages 82–90. 2017.
- [44] V. CHALANA, D.T. LINKER, D.R. HAYNOR et Yongmin KIM : A multiple active contour model for cardiac boundary detection on echocardiographic sequences. *IEEE Transactions on Medical Imaging*, 15(3):290–298, Jun 1996. ISSN 0278-0062.
- [45] T. CHEN, J. BABB, P. KELLMAN, L. AXEL et D. KIM : Semiautomated segmentation of myocardial contours for fast strain analysis in cine displacement-encoded MRI. *IEEE Trans. Med. Imaging*, 27(8):1084–1094, Aug 2008. ISSN 0278-0062.
- [46] T. CHEN, X. WANG, S. CHUNG, D. METAXAS et L. AXEL : Automated 3d motion tracking using gabor filter bank, robust point matching, and deformable models. *Medical Imaging, IEEE Transactions on*, 29:1 – 11, 02 2010.
- [47] X. CHEN, H. XIE, R. ERKAMP, K. KIM, C. JIA, R. M RUBIN et M. O'DONNELL : 3-d correlation-based speckle tracking. *Ultrasonic imaging*, 27:21–36, 02 2005.
- [48] J. CHENG et J.Y. LU : Extended high-frame rate imaging method with limited-diffraction beams. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 53(5): 880–899, 2006.
- [49] S. CHENG, M. G. LARSON, E. L. MCCABE, E. OSYPIUK, B. T. LEHMAN, P. STANCHEV, J. ARAGAM, E. J. BENJAMIN, S. D. SOLOMON et R. S. VASAN : Reproducibility of speckle-tracking-based strain measures of left ventricular function in a community-based study. *Journal of the American Society of Echocardiography*, 26(11):1258–1266.e2, Nov 2013. ISSN 0894-7317. URL <https://doi.org/10.1016/j.echo.2013.07.002>.
- [50] P. CLARYSSE, J. TAFAZZOLI, P. DELACHARTRE et P. CROISILLE : Simulation based evaluation of cardiac motion estimation methods in tagged-mr image sequences. *Journal of Cardiovascular Magnetic Resonance*, 13(1):1–1, 2011.
- [51] T.F. COOTES, C.J. TAYLOR, D.H. COOPER et J. GRAHAM : Active shape models-their training and application. *Computer Vision and Image Understanding*, 61(1):38 – 59, 1995. ISSN 1077-3142.
- [52] M. CORREIA, J. PROVOST, M. TANTER et M. PERNOT : 4d ultrafast ultrasound flow imaging :in vivo quantification of arterial volumetric flow rate in a single heartbeat. *Physics in Medicine and Biology*, 61(23):L48–L61, nov 2016.
- [53] C. CORSI, G. SARACINO, A. SARTI et C. LAMBERTI : Left ventricular volume estimation for real-time three-dimensional echocardiography. *IEEE Transactions on Medical Imaging*, 21(9):1202–1208, 2002. ISSN 0278-0062.
- [54] M. De CRAENE, S. MARCHESSEAU, B. HEYDE, H. GAO, M. ALESSANDRINI, O. BERNARD, G. PIELLA, A. R. PORRAS, L. TAUTZ, A. HENNEMUTH, A. PRAKOSA, H. LIEBGOTT,

- O. SOMPHONE, P. ALLAIN, S. Makram EBEID, H. DELINGETTE, M. SERMESANT, J. D'HOOGHE et E. SALOUX : 3d strain assessment in ultrasound (straus) : A synthetic comparison of five tracking methodologies. *IEEE Transactions on Medical Imaging*, 32 (9):1632–1646, Sept 2013. ISSN 0278-0062.
- [55] M. De CRAENE, G. PIELLA, O. CAMARA, N. DUCHATEAU, E. SILVA, A. DOLTRA, J. D'HOOGHE, J. BRUGADA, M. SITGES et A. F. FRANGI : Temporal diffeomorphic free-form deformation : Application to motion and strain estimation from 3d echocardiography. *Medical Image Analysis*, 16(2):427–450, 2012.
- [56] J. CROSBY, B. H. AMUNDSEN, T. HERGUM, E. W. REMME, S. LANGELAND et H. TORP : 3-d speckle tracking for assessment of regional left ventricular function. *Ultrasound in Medicine and Biology*, 35(3):458–471, Mar 2009.
- [57] W. R. CRUM, E. BERRY, J. P. RIDGWAY, U. M. SIVANANTHAN, L. B. TAN et M. A. SMITH : Simulation of two-dimensional tagged mri. *Journal of Magnetic Resonance Imaging*, 7 (2):416–424, 1997.
- [58] W. R. CRUM, E. BERRY, J. P. RIDGWAY, U. M. SIVANANTHAN, L. B. TAN et M. A. SMITH : Frequency-domain simulation of mr tagging. *Journal of Magnetic Resonance Imaging*, 8 (5):1040–1050, 1998.
- [59] L. DEMI, A. RAMALLI, G. GIANNINI et M. MISCHI : In vitro and in vivo tissue harmonic images obtained with parallel transmit beamforming by means of orthogonal frequency division multiplexing [correspondence]. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 62(1):230–235, January 2015.
- [60] L. DEMI, M. D. VERWEIJ et K. W. a. VAN DONGEN : Parallel transmit beamforming using orthogonal frequency division multiplexing applied to harmonic imaging—a feasibility study. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 59(11):2439–2447, November 2012.
- [61] B. DENARIE, T. BJASTAD et H. TORP : Multi-line transmission in 3-d with reduced crosstalk artifacts : a proof of concept study. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 60(8):1708–1718, August 2013.
- [62] B. DENARIE, T. A. TANGEN, I. K. EKROLL, N. ROLIM, H. TORP, T. BJASTAD et L. LOVSTAKKEN : Coherent plane wave compounding for very high frame rate ultrasonography of rapidly moving targets. *Medical Imaging, IEEE Transactions on*, 32 (7):1265–1276, 2013.
- [63] B. DENARIE, T. A. TANGEN, I. K. EKROLL, N. ROLIM, H. TORP, T. BJASTAD et L. LOVSTAKKEN : Coherent plane wave compounding for very high frame rate ultrasonography of rapidly moving targets. *IEEE Transactions on Medical Imaging*, 32 (7):1265–1276, July 2013.
- [64] L. R. DICE : Measures of the amount of ecologic association between species. *Ecology*, 26 (3):pp. 297–302, 1945. ISSN 00129658.
- [65] E. DIKICI et F. ORDERUD : Generalized step criterion edge detectors for kalman filter based left ventricle tracking in 3D+T echocardiography. In *Statistical Atlases and Computational Models of the Heart. Imaging and Modelling Challenges*, pages 261–269, 2012. ISBN 978-3-642-36960-5.

-
- [66] P. DOLLÁR et C. L. ZITNICK : Fast Edge Detection Using Structured Forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(8):1558–1570, 2015.
- [67] J. DOMINGOS, R. STEBBING et J. NOBLE : Endocardial segmentation using structured random forests in 3D echocardiography. In *Proc. MICCAI Challenge on Echocardiographic Three-Dimensional Ultrasound Segmentation (CETUS), Boston, MIDAS Journal*, pages 33–40. 2014.
- [68] A. DRUKAREV, K. KONSTANTINIDES et G. SEROUSSI : Beam transformation techniques for ultrasonic medical imaging. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 40(6):717–726, Nov 1993.
- [69] Q. DUAN, E. ANGELINI, S. HERZ, C. INGRASSIA, K. COSTA, J. HOLMES, S. HOMMA et A. LAINE : Region-based endocardium tracking on real-time three-dimensional ultrasound. *Ultrasound In Medicine and Biology*, 35(2):256–265, 2009.
- [70] Q. DUAN, E. D. ANGELINI, O. GÉRARD, S. HOMMA et A. F. LAINE : Comparing optical-flow based methods for quantification of myocardial deformations on rt3d ultrasound. *3rd IEEE International Symposium on Biomedical Imaging : Nano to Macro, 2006.*, pages 173–176, 2006.
- [71] Q. DUAN, E. D. ANGELINI et A. F. LAINE : Real-time segmentation by active geometric functions. *Computer methods and programs in biomedicine*, 98(3):223–230, 2010.
- [72] Q. DUAN, P. MOIREAU, E. D. ANGELINI, D. CHAPELLE et A. F. LAINE : Simulation of 3d ultrasound with a realistic electro-mechanical model of the heart. In Frank B. SACHSE et Gunnar SEEMANN, éditeurs : *Functional Imaging and Modeling of the Heart*, pages 463–473, Berlin, Heidelberg, 2007. Springer Berlin Heidelberg.
- [73] D. T. DUBBERSTEIN et O. T. VON RAMM : Methods and systems for ultrasound scanning using spatially and spectrally separated transmit ultrasound beams, 2000, US Patent.
- [74] D. DUBOIS et E. F. DUBOIS : A formula to estimate the approximate surface area if height and weight be known. *archives of internal medicine.*, *Archives of Internal Medicine*, 17:863–871, 1916.
- [75] N. DUCHATEAU, M. De CRAENE, P. ALLAIN, E. SALOUX et M. SERMESANT : Infarct localization from myocardial deformation : Prediction and uncertainty quantification by regression from a low-dimensional space. *IEEE transactions on medical imaging*, 35(10): 2340–2352, 2016.
- [76] S. Makram EBEID et O. SOMPHONE : Non-rigid image registration using a hierarchical partition of unity finite element method. pages 1–8, 10 2007.
- [77] I. K. EKROLL, T. DAHL, H. TORP et L. LØVSTAKKEN : Combined vector velocity and spectral doppler imaging for improved imaging of complex blood flow in the carotid arteries. *Ultrasound in Medicine and Biology*, 40(7):1629–1640, 2014.
- [78] I. K. EKROLL, M. M. VOORMOLEN, O. K. V. STANDAL, J. M. RAU et L. LOVSTAKKEN : Coherent compounding in doppler imaging. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 62(9):1634–1643, Sept 2015.

- [79] A. ELEN, H. F. CHOI, D. LOECKX, H. GAO, P. CLAUS, P. SUETENS, F. MAES et J. D'HOOGHE : Three-dimensional cardiac strain estimation using spatio-temporal elastic registration of ultrasound images : A feasibility study. *IEEE Transactions on Medical Imaging*, 27(11):1580–1591, Nov 2008. ISSN 0278-0062.
- [80] P. M. ELLIOTT, A. ANASTASAKIS, M. A. BORGER, M. BORGGREFE, F. CECCHI, P. CHARRON, A. A. HAGEGE, A. LAFONT, G. LIMONGELLI *et al.* : 2014 ESC guidelines on diagnosis and management of hypertrophic cardiomyopathy the task force for the diagnosis and management of hypertrophic cardiomyopathy of the european society of cardiology (ESC). *Eur. Heart J.*, 35(39):2733–2779, 2014.
- [81] O. EMAD, I. A. YASSINE et A. S. FAHMY : Automatic localization of the left ventricle in cardiac mri images using deep learning. *In International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 683–686, Aug 2015.
- [82] B. Wenjia et AL. : Human-level cmr image analysis with deep fully convolutional networks. *arXiv preprint :1710.09289*, 2017.
- [83] J. FAURIE, M. BAUDET, J. PORÉE, G. CLOUTIER, F. TOURNOUX et D. GARCIA : Coupling myocardium and vortex dynamics in diverging-wave echocardiography. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, pages 1–1, 2018.
- [84] J. FAURIE, M. BAUDET, J. PORÉE, G. CLOUTIER, F. TOURNOUX et D. GARCIA : Coupling myocardium and vortex dynamics in diverging-wave echocardiography. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 66(3):425–432, 2019.
- [85] S. E. FISCHER, G. C. MCKINNON, S. E. MAIER et P. BOESIGER : Improved myocardial tagging contrast. *Magnetic resonance in medicine*, 30(2):191–200, 1993.
- [86] E. D. FOLLAND, A. F. PARISI, P. F. MOYNIHAN, D. R. JONES, C. L. FELDMAN et D. E. TOW : Assessment of left ventricular ejection fraction and volumes by real-time, two-dimensional echocardiography. A comparison of cineangiographic and radionuclide techniques. *Circulation*, 60(4):760–766, 1979.
- [87] K. L. GAMMELMARK et J. A. JENSEN : 2-d tissue motion compensation of synthetic transmit aperture images. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 61(4):594–610, 2014.
- [88] H. GAO, H. F. CHOI, P. CLAUS, S. BOONEN, S. JAECQUES, G. H. Van LENTHE, G. Van Der PERRE, W. LAURIKS et J. D'HOOGHE : A fast convolution-based methodology to simulate 2-d/3-d cardiac ultrasound images. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 56(2):404–409, February 2009. ISSN 0885-3010.
- [89] H. GAO, T. HERGUM, H. TORP et J. D'HOOGHE : Comparison of the performance of different tools for fast simulation of ultrasound data. *Ultrasonics*, 52(5):573 – 577, 2012. ISSN 0041-624X. URL <http://www.sciencedirect.com/science/article/pii/S0041624X12000248>.
- [90] D. GARCIA, L.L. TARNEC, S. MUTH, E. MONTAGNON, J. PORÉE et G. CLOUTIER : Stolt's f-k migration for plane wave ultrasound imaging. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 60(9):1853–1867, 2013.

-
- [91] J.B. GARRISON, W.L. EBERT, R.E. JENKINS, S.M. YIONOULIS, H. MALCOM, G.A. HEYLER, A.A. SHOUKAS, W.L. MAUGHAN et K. SAGAWA : Measurement of three-dimensional positions and motions of large numbers of spherical radiopaque markers from biplane cineradiograms. *Computers and Biomedical Research*, 15(1):76 – 96, 1982.
- [92] I. GAZDAG et P. SGUAZZERO : Migration of seismic data. *Proceedings of the IEEE*, 72(10):1302–1315, 1984.
- [93] J.L. GENNISSON, T. DEFFIEUX, E. MACÉ, G. MONTALDO, M. FINK et M. TANTER : Viscoelastic and anisotropic mechanical properties of in vivo muscle tissue assessed by supersonic shear imaging. *Ultrasound in medicine and biology*, 36(5):789–801, 2010.
- [94] T. GLATARD, C. LARTIZIEN, B. GIBAUD, R. F. Da SILVA, G. FORESTIER, F. CERVENANSKY, M. ALESSANDRINI, H. BENOIT-CATTIN, O. BERNARD, S. CAMARASU-POP *et al.* : A virtual imaging platform for multi-modality medical image simulation. *IEEE Transactions on Medical Imaging*, 32(1):110–118, 2013.
- [95] D. GROSGEORGE, C. PETITJEAN, J.-N. DACHER et S. RUAN : Graph cut segmentation with a statistical shape model in cardiac MRI. *Comput. Vis. Image Underst.*, 117(9):1027 – 1035, 2013.
- [96] A. GROTH, J. WEESE et H. LEHMANN : Robust left ventricular myocardium segmentation for multi-protocol mr. *In SPIE Medical Imaging*. International Society for Optics and Photonics, 2012.
- [97] A. HAASE : Snapshot flash mri. applications to t1, t2, and chemical-shift imaging. *Magnetic Resonance in Medicine*, 13(1):77–89, 1990.
- [98] R. HADDAD, P. CLARYSSE, M. ORKISZ, P. CROISILLE, D. REVEL et I. E. MAGNIN : A realistic anthropomorphic numerical model of the beating heart. *In International Workshop on Functional Imaging and Modeling of the Heart*, pages 384–393. Springer, 2005.
- [99] Rana HADDAD, Isabelle E MAGNIN et P CLARYSSE : A new fully-digital anthropomorphic and dynamic thorax/heart model. *In Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pages 5999–6002. IEEE, 2007.
- [100] J. HANSEGARD, F. ORDERUD et S. RABBEN : Real-time active shape models for segmentation of 3D cardiac ultrasound. *In Computer Analysis of Images and Patterns*, pages 157–164, 2007. ISBN 978-3-540-74271-5.
- [101] H. HASEGAWA et H. KANAI : High-frame-rate echocardiography using diverging transmit beams and parallel receive beamforming. *Journal of Medical Ultrasonics*, 38(3):129–140, 2011.
- [102] M. HAVAEI, N. GUIZARD, H. LAROCHELLE et P-M JODOIN : Deep learning trends for focal brain pathology segmentation in mri. *In book Machine Learning for Health Informatics, State-of-the-art and future challenges*, pages 125–148, LNAI 9605, Springer, 2015.
- [103] T. HERGUM, T. BJASTAD, K. KRISTOFFERSEN et H. TORP : Parallel beamforming using synthetic transmit beams. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 54(2):271–280, February 2007.

- [104] B. HEYDE, M. ALESSANDRINI, J. HERMANS, D. BARBOSA, P. CLAUS et J. D'HOOGHE : Anatomical image registration using volume conservation to assess cardiac deformation from 3d ultrasound recordings. *IEEE transactions on medical imaging*, 35(2):501–511, 2016.
- [105] B. HEYDE, D. BARBOSA, P. CLAUS, F. MAES et J. D'HOOGHE : Three-dimensional cardiac motion estimation based on non-rigid image registration using a novel transformation model adapted to the heart. *In STACOM, Springer*, pages 142–150, 2013.
- [106] K. HOR, W. M. GOTTLIEBSON, C. CARSON, E. WASH, J. CNOTA, R. FLECK, J. WANSAPURA, P. KLIMECZEK, H. R. AL-KHALIDI, E. CHUNG, W. BENSON et W. MAZUR : Comparison of magnetic resonance feature tracking for strain calculation with harmonic phase imaging analysis. *JACC. Cardiovascular imaging*, 3:144–51, 02 2010.
- [107] Kan HOR, Rolf BAUMANN, Gianni PEDRIZZETTI, G TONTI, William M GOTTLIEBSON, Michael TAYLOR, Woody BENSON et Wojciech MAZUR : Magnetic resonance derived myocardial strain assessment using feature tracking. *Journal of visualized experiments : JoVE*, 48, 02 2011.
- [108] J. HSIEH et SPIE (SOCIETY) : *Computed tomography : principles, design, artifacts, and recent advances*. Bellingham, Wash. SPIE Hoboken, N. J. John Wiley & Sons, 2009.
- [109] D. P. HUTTENLOCHER, G. A. KLANDERMAN et W. J. RUCKLIDGE : Comparing images using the Hausdorff distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(9):850–863, Sep 1993. ISSN 0162-8828.
- [110] E. ILG, N. MAYER, T. SAIKIA, M. KEUPER, A. DOSOVITSKIY et T. BROX : FlowNet 2.0 : Evolution of optical flow estimation with deep networks. *In 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1647–1655, 2017.
- [111] F. ISENSEE, P. JAEGER, P. FULL, I. WOLF, S. ENGELHARDT et K. H. MAIER-HEIN : Automatic cardiac disease assessment on cine-mri via time-series segmentation and domain specific features. *In Proc. STACOM-MICCAI, LNCS, volume 10663*, pages 120–129, 2017.
- [112] J. I. JACKSON, C. H. MEYER, D. G. NISHIMURA et A. MACOVSKI : Selection of a convolution function for fourier inversion using gridding [computerised tomography application]. *IEEE Transactions on Medical Imaging*, 10(3):473–478, 1991.
- [113] Y. JANG, S. HA, S. KIM, Y. HONG et H-J. CHANG : Automatic segmentation of lv and rv in cardiac mri. *In Proc. STACOM-MICCAI, LNCS, volume 10663*, pages 161–169, 2017.
- [114] C. JENKINS, S. MOIR, J. CHAN, D. RAKHIT, B. HALUSKA et T. H. MARWICK : Left ventricular volume measurement with echocardiography : a comparison of left ventricular opacification, three-dimensional echocardiography, or both with magnetic resonance imaging. *European Heart Journal*, 30(1):98–106, 2009. URL <http://dx.doi.org/10.1093/eurheartj/ehn484>.
- [115] J. A. JENSEN : Field : A program for simulating ultrasound systems. *In 10th Nordicbaldic Conf. Biomed. Imag.*, pages 351–353, 1996.
- [116] J. A. JENSEN : Field : A program for simulating ultrasound systems. *In 10th Nordicbaldic Conf. Biomed. Imag.*, pages 351–353, 1996.

-
- [117] J. A. JENSEN et N. B. SVENDSEN : Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 39(2):262–267, March 1992. ISSN 0885-3010.
- [118] J. A. JENSEN et N. B. SVENDSEN : Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 39(2):262–267, March 1992.
- [119] J.A. JENSEN, S.I. NIKOLOV, K.L. GAMMELMARK et M.H. PEDERSEN : Synthetic aperture ultrasound imaging. *Ultrasonics*, 44:e5–e15, 2006.
- [120] J.A. JENSEN et N.B. SVENDSEN : Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 39(2):262–267, 1992.
- [121] T. H. JOCHIMSEN, A. SCHÄFER, R. BAMMER et M. E. MOSELEY : Efficient simulation of magnetic resonance imaging with bloch–torrey equations using intra-voxel magnetization gradients. *Journal of Magnetic Resonance*, 180(1):29–38, 2006.
- [122] I. JOLLIFFE : *Principal Component Analysis*. John Wiley & Sons, Ltd, 2014. ISBN 9781118445112.
- [123] K. H. JSHEIKH, S. W. SMITH, O. V. RAMM et J. KISSLO : Real-time, three-dimensional echocardiography : Feasibility and initial us. *Echocardiography*, 8(1):119 – 125, 1991.
- [124] R. JUANG, E.R. MCVEIGH, B. HOFFMANN, D. YUH et P. BURLINA : Automatic segmentation of the left-ventricular cavity and atrium in 3D ultrasound using graph cuts and the radial symmetry transform. *In IEEE International Symposium on Biomedical Imaging*, pages 606–609, 2011.
- [125] H. B. KAUSE, O. FILATOVA, R. DUIJS, M. BRUURMIJN, A. FUSTER, J. WESTENBERG, L. FLORACK et H. van ASSEN : Direct myocardial strain assessment from frequency estimation in tagging mri. 01 2014.
- [126] M. KHENED, V. ALEX et G. KRISHNAMURTHI : Densely connected fully convolutional network for short-axis cardiac cine mr image segmentation and heart diagnosis using random forest. *In Proc. STACOM-MICCAI, LNCS, volume 10663*, pages 140–151, 2017.
- [127] K. S. KIM, J. S. HWANG, J. S. JEONG et T. K. SONG : An efficient motion estimation and compensation method for ultrasound synthetic aperture imaging. *Ultrasonic imaging*, 24(2):81–99, 2002.
- [128] D. P. KINGMA et M. WELLING : Auto-encoding variational bayes. *In International Conference on Learning Representations*, 2013.
- [129] S. A. KLEIJN, M. F. A. ALY, C. B. TERWEE, A. C. van ROSSUM et O. KAMP : Reliability of left ventricular volumes and function measurements using three-dimensional speckle tracking echocardiography. *European Heart Journal - Cardiovascular Imaging*, 13(2):159–168, 2012. URL <http://dx.doi.org/10.1093/ejechocard/jer174>.
- [130] B. KONG, Y. ZHAN, M. SHIN, T. DENNY et S. ZHANG : Recognizing end-diastole and end-systole frames via deep temporal regression network. *In Proc. MICCAI*, pages 264–272. 2016.

- [131] C. KRITTANAWONG, S. M. NARAYAN, J. L. HALPERIN, R. S. ROSENSON, U. BABER, K. W. JOHNSON, Z. WANG, M. AYDAR, J. K. MIN et W. H. W. TANG : Deep learning for cardiovascular medicine : a practical primer. *European Heart Journal*, 02 2019.
- [132] P. KRUIZINGA, F. MASTIK, N. de JONG, A. van der STEEN et G. van SOEST : Plane-wave ultrasound beamforming using a nonuniform fast fourier transform. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 59(12):2684–2691, 2012.
- [133] J. P. A. KUIJER, E. JANSEN, J. Tim MARCUS, A. C. van ROSSUM et R. M. HEETHAAR : Improved harmonic phase myocardial strain maps. *Magnetic Resonance in Medicine*, 46: 993–999, 11 2001.
- [134] A. LALANDE, M. GARREAU et F. FROUIN : Evaluation of cardiac structure segmentation in cine magnetic resonance imaging. In *Multi-modality Cardiac Imaging : Processing and Analysis*, pages 171–215. Iste, 2015.
- [135] R. M. LANG, L. P. BADANO, V. MOR-AVI, J. AFILALO *et al.* : Recommendations for Cardiac Chamber Quantification by Echocardiography in Adults : An Update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. *Circulation*, 28(1):1–39, 2015.
- [136] R. M. LANG, L. P. BADANO, W. TSANG, D. H. ADAMS, E. AGRICOLA, T. BUCK, F. F. FALETRA, A. FRANKE, J. HUNG, L. P. de ISLA, O. KAMP, J. D. KASPRZAK, P. LANCELLOTTI, T. H. MARWICK, M. L. MCCULLOCH, M. J. MONAGHAN, P. NIHOYANNOPOULOS, N. G. PANDIAN, P. A. PELLIKKA, M. PEPI, D. A. ROBERSON, S. K. SHERNAN, G. S. SHIRALI, L. SUGENG, F. J. TEN CATE, M. A. VANNAN, J. L. ZAMORANO et W. A. ZOGHBI : EAE/ASE recommendations for image acquisition and display using three-dimensional echocardiography. *European Heart Journal - Cardiovascular Imaging*, 13(1):1–46, Jan 2012.
- [137] S. LECLERC, T. GRENIER, F. ESPINOSA et O. BERNARD : A fully automatic and multi-structural segmentation of the left ventricle and the myocardium on highly heterogeneous 2D echocardiographic data. In *IEEE International Ultrasonics Symposium (IUS)*, 2017.
- [138] S. LECLERC, E. SMISTAD, T. GRENIER, C. LARTIZIEN, A. OSTVIK, F. ESPINOSA, P. M. JODOIN, L. LOVSTAKKEN et O. BERNARD : Deep learning applied to multi-structures segmentation in 2D echocardiography : a preliminary investigation of the required database size. In *2018 IEEE International Ultrasonics Symposium (IUS)*, pages 1–1, Oct 2018.
- [139] S. LECLERC, E. SMISTAD, J. PEDROSA, A. ØSTVIK, F. CERVENANSKY, F. ESPINOSA, T. ESPELAND, E. A. R. BERG, P. JODOIN, T. GRENIER, C. LARTIZIEN, J. D’HOOGHE, L. LOVSTAKKEN et **O. Bernard** : Deep learning for segmentation using an open large-scale dataset in 2d echocardiography. *IEEE Transactions on Medical Imaging*, pages 1–1, 2019.
- [140] M. J. LEDESMA-CARBAYO, J. KYBIC, M. DESCO, A. SANTOS, M. SUHLING, P. HUNZIKER et M. UNSER : Spatio-temporal nonrigid registration for ultrasound cardiac motion estimation. *IEEE Transactions on Medical Imaging*, 24(9):1113–1126, Sept 2005. ISSN 0278-0062.

-
- [141] V. LEMPITSKY, M. VERHOEK, J. A. NOBLE et A. BLAKE : Random forest classification for automatic delineation of myocardium in real-time 3d echocardiography. *In Functional Imaging and Modeling of the Heart*, pages 447–456, 2009. ISBN 978-3-642-01931-9.
- [142] M. LENGE, A. RAMALLI, E. BONI, H. LIEBGOTT, C. CACHARD et P. TORTOLI : High-frame-rate 2-d vector blood flow imaging in the frequency domain. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 61(9):1504–1514, 2014.
- [143] K. Y. LEUNG, M. DANILOUCHKINE, M. van STRALEN, N. de JONG, A. van der STEEN et J. G. BOSCH : Left ventricular border tracking using cardiac motion models and optical flow. *Ultrasound In Medicine and Biology*, 37(4):605–616, 2011.
- [144] K. Y. LEUNG, M. van STRALEN, G. van BURKEN, N. de JONG et J. G. BOSCH : Automatic active appearance model segmentation of 3D echocardiograms. *In IEEE International Symposium on Biomedical Imaging*, pages 320–323, 2010.
- [145] K. Y. E. LEUNG et J. G. BOSCH : Automated border detection in three-dimensional echocardiography : principles and promises. *European Journal of Echocardiography*, 11(2):97–108, 2010.
- [146] B. LI, Y. LIU, J. O. CHRISTOPHER, R. C. BRETT et A. Y. ALISTAIR : In-line automated tracking for ventricular function with magnetic resonance imaging. *JACC. Cardiovasc. Imaging*, 3 8:860–6, 2010.
- [147] G. LITJENS, T. KOOI, B. E. BEJNORDI, A. A. A. SETIO, F. CIOMPI, M. GHAFORIAN, J. A. W. M. van der LAAK, B. van GINNEKEN et C. I. SÁNCHEZ : A survey on deep learning in medical image analysis. *Med. Image Anal.*, 42(Supplement C):60 – 88, 2017. ISSN 1361-8415.
- [148] H. LIU, H. HU, X. XU et E. SONG : Automatic left ventricle segmentation in cardiac MRI using topological stable-state thresholding and region restricted dynamic programming. *Acad. Radiol.*, 19(6):723–731, 2012. ISSN 1076-6332.
- [149] X. LIU, K. Z. ABD-ELMONIEM, M. STONE, E. Z. MURANO, J. ZHUO, R. P. GULLAPALLI et J. L. PRINCE : Incompressible deformation estimation algorithm (idea) from tagged mr images. *IEEE Transactions on Medical Imaging*, 31(2):326–340, 2012.
- [150] X. LIU et J. PRINCE : Shortest path refinement for motion estimation from tagged mr images. *IEEE transactions on medical imaging*, 29:1560–72, 03 2010.
- [151] J. LONG, E. SHELFHAMER et T. DARRELL : Fully convolutional networks for semantic segmentation. *In Proc. CVPR*, pages 3431–3440, 2014.
- [152] R. LOPATA, M. M. NILLESEN, J. THIJSSSEN, L. KAPUSTA et C. de KORTE : Three-dimensional cardiac strain imaging in healthy children using rf-data. *Ultrasound in medicine and biology*, 37:1399–408, 09 2011.
- [153] C. H. LORENZ, E. S. WALKER, V. L. MORGAN, S. S. KLEIN et T. P. GRAHAM : Normal human right and left ventricular mass, systolic function, and gender differences by cine magnetic resonance imaging. *J Cardiovasc Magn Reson*, 1(1):1097–6647, 1999.

- [154] J.Y. LU : 2d and 3d high frame rate imaging with limited diffraction beams. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 44(4):839–856, July 1997.
- [155] J.Y. LU : Experimental study of high frame rate imaging with limited diffraction beams. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 45(1):84–97, 1998.
- [156] A. MAKHZANI, J. SHLENS, N. JAITLY et I. GOODFELLOW : Adversarial autoencoders. *In International Conference on Learning Representations*, 2016.
- [157] R. MALLART et M. FINK : Improved imaging rate through simultaneous transmission of several ultrasound beams. *In International Society for Optics and Photonics*, pages 120–130, 1992.
- [158] S. MARCHESSEAU, H. DELINGETTE, M. SERMESANT et N. AYACHE : Fast parameter calibration of a cardiac electromechanical model from medical images based on the unscented transform. *Biomechanics and modeling in mechanobiology*, 12(4):815–831, 2013.
- [159] S. MARCHESSEAU, H. DELINGETTE, M. SERMESANT, M. SORINE, K. RHODE, S.G. DUCKETT, C.A. RINALDI, R. RAZAVI et N. AYACHE : Preliminary specificity study of the bestel–clément–sorine electromechanical model of the heart using parameter calibration from medical images. *Journal of the Mechanical Behavior of Biomedical Materials*, 20:259 – 271, 2013. ISSN 1751-6161. URL <http://www.sciencedirect.com/science/article/pii/S1751616112003335>.
- [160] F. I. MARCUS, W. J. MCKENNA, D. SHERRILL, C. BASSO, B. BAUCE, D. A. BLUEMKE, H. CALKINS, D. CORRADO, M. G.P.J. COX *et al.* : Diagnosis of arrhythmogenic right ventricular cardiomyopathy / dysplasia. *Circulation*, 121(13):1533–1541, 2010.
- [161] E. MARET, T. TODT, L. BRUDIN, E. NYLANDER, E. SWAHN, J. L. OHLSSON et J. ENGVALL : Functional measurement based on feature tracking of cine magnetic resonance images identify left ventricular segments with myocardial scar. *Cardiovascular ultrasound*, 7:53, 11 2009.
- [162] R. MEHTA et J. SIVASWAMY : M-net : A convolutional neural network for deep brain structure segmentation. *In Proc. ISBI*, pages 437–440, 2017.
- [163] S. MENDIS : Global status report on noncommunicable diseases. Rapport technique, World Health Organization reports, 2014.
- [164] Ch. A. MILLER, P. JORDAN, A. BORG, R. ARGYLE, D. CLARK, K. PEARCE et M. SCHMITT : Quantification of left ventricular indices from SSFP cine imaging : Impact of real-world variability in analysis methodology and utility of geometric modeling. *J. Magn. Reson. Imaging*, 37(5):1213–1222, 2013. ISSN 1522-2586.
- [165] S. C. MITCHELL, J. G. BOSCH, B. P. F. LELIEVELDT, R. J. van der GEEST, J. H. C. REIBER et M. SONKA : 3-d active appearance models : segmentation of cardiac mr and ultrasound images. *IEEE Trans. Med. Imaging*, 21(9):1167–1178, 2002.
- [166] S. MONDILLO, M. GALDERISI, D. MELE, M. CAMELI, V. S. LOMORIELLO, V. ZACÀ, P. BALLO, A. D’ANDREA, D. MURARU, M. LOSI, E. AGRICOLA, A. D’ERRICO,

-
- S. BURALLI, S. SCIOMER, S. NISTRI et L. BADANO : Speckle-tracking echocardiography. *Journal of Ultrasound in Medicine*, 30(1):71–83, 2011.
- [167] G. MONTALDO, M. TANTER, J. BERCOFF, N. BENECH et M. FINK : Coherent plane-wave compounding for very high frame rate ultrasonography and transient elastography. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 56(3):489–506, 2009.
- [168] William MOODY, Robin J. TAYLOR, Nicola EDWARDS, Colin CHUE, Fraz UMAR, Tiffany J. TAYLOR, Charles FERRO, Alistair YOUNG, Jonathan N. TOWNEND, Francisco LEYVA et Richard STEEDS : Comparison of magnetic resonance feature tracking for systolic and diastolic strain and strain rate calculation with spatial modulation of magnetization imaging analysis. *Journal of Magnetic Resonance Imaging*, 41, 03 2014.
- [169] V. MOR-AVI, C. JENKINS, H. KUHL, H. NESSER, T. MARWICK, A. FRANKE, C. EBNER, B. FREED, R. STERINGER-MASCHERBAUER, H. POLLARD, L. WEINERT, J. NIEL, L. SUGENG et R. LANG : Real-time 3-dimensional echocardiographic quantification of left ventricular volumes : Multicenter study for validation with magnetic resonance imaging and investigation of sources of error. *Journal of the American College of Cardiology*, 1 (4):413–423, 2008. ISSN 1936-878X.
- [170] V. MOR-AVI, R. M. LANG, L. P. BADANO, M. BELOHLAVEK, N. M. CARDIM, G. DERUMEAUX, M. GALDERISI, T. MARWICK, S. F. NAGUEH, P. P. SENGUPTA, R. SICARI, O. A. SMISETH, B. SMULEVITZ, M. TAKEUCHI, J. D. THOMAS, M. VANNAN, J. VOIGT et J. L. ZAMORANO : Current and evolving echocardiographic techniques for the quantitative evaluation of cardiac mechanics : Ase/ae consensus statement on methodology and indications : Endorsed by the japanese society of echocardiography. *Journal of the American Society of Echocardiography*, 24(3):277–313, 2011.
- [171] M. MULLER, D. AIT-BELKACEM, M. HESSABI, J. L. GENNISSON, G. GRANGÉ, F. GOFFINET, E. LECARPENTIER, D. CABROL, M. TANTER et V. TSATSARIS : Assessment of the cervix in pregnant women using shear wave elastography : A feasibility study. *Ultrasound in Medicine and Biology*, 41(11):2789–2797, 2015.
- [172] D. MURARU, L. BADANO, G. PICCOLI, P. GIANFAGNA, L. DEL MESTRE, D. ERMACORA et A. PROCLEMER : Validation of a novel automated border-detection algorithm for rapid and accurate quantitation of left ventricular volumes based on three-dimensional echocardiography. *European journal of echocardiography*, 11(4):359–368, 2010.
- [173] A. MYRONENKO, X. SONG et D. J. SAHN : Maximum likelihood motion estimation in 3d echocardiography through non-rigid registration in spherical coordinates. In Nicholas AYACHE, Hervé DELINGETTE et Maxime SERMESANT, éditeurs : *Functional Imaging and Modeling of the Heart*, pages 427–436, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg.
- [174] A. NEWELL, K. YANG et J. DENG : Stacked hourglass networks for human pose estimation. In Bastian LEIBE, Jiri MATAS, Nicu SEBE et Max WELLING, éditeurs : *Computer Vision – ECCV 2016*, pages 483–499, Cham, 2016. Springer International Publishing. ISBN 978-3-319-46484-8.
- [175] T. A. NGO, Z. LU et G. CARNEIRO : Combining deep learning and level set for the automated segmentation of the left ventricle of the heart from cardiac cine magnetic resonance. *Med. Image Anal.*, 35:159–171, 2017.

- [176] L. NING et J. S. DUNCAN : Generalized robust point matching using an extended free-form deformation model : application to cardiac images. *In 2004 2nd IEEE International Symposium on Biomedical Imaging : Nano to Macro (IEEE Cat No. 04EX821)*, pages 320–323 Vol. 1, 2004.
- [177] J. A. NOBLE et D. BOUKERROUI : Ultrasound image segmentation : a survey. *IEEE Transactions on Medical Imaging*, 25(8):987–1010, 2006.
- [178] R.M. NORRIS, H.D. WHITE, D.B. CROSS, C.J. WILD et R.M. WHITLOCK : Prognosis after recovery from myocardial infarction : the relative importance of cardiac dilatation and coronary stenoses. *Eur. Heart J.*, 13:1611–1618, 1992.
- [179] Ö. ÇIÇEK, A. ABDULKADIR, S. S. LIENKAMP, T. BROX et O. RONNEBERGER : 3D U-Net : Learning Dense Volumetric Segmentation from Sparse Annotation. *In Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*, pages 424–432, Cham, 2016. Springer International Publishing.
- [180] N. ODDERSHEDE et J. A. JENSEN : Effects influencing focusing in synthetic aperture vector flow imaging. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, 54(9):1811–1825, 2007.
- [181] M. O’DONNELL, A. R. SKOVORODA, B. M. SHAPO et S. Y. EMELIANOV : Internal displacement and strain imaging using ultrasonic speckle tracking. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 41(3):314–325, 1994.
- [182] O. OKTAY, W. BAI, M. LEE, R. GUERRERO, K. KAMNITSAS, J. CABALLERO, A. de MARVAO, S. COOK, D. O’REGAN et D. RUECKERT : Multi-input cardiac image super-resolution using convolutional neural networks. *In Proc. MICCAI*, pages 246–254. 2016.
- [183] O. OKTAY, E. FERRANTE, K. KAMNITSAS, M. HEINRICH, W. BAI, J. CABALLERO, S. A. COOK, A. de MARVAO, T. DAWES, D. P. O’REGAN, B. KAINZ, B. GLOCKER et D. RUECKERT : Anatomically Constrained Neural Networks (ACNNs) : Application to Cardiac Image Enhancement and Segmentation. *IEEE Transactions on Medical Imaging*, 37(2):384–395, 2018.
- [184] O. OKTAY, A. GOMEZ, K. KERAUDREN, A. SCHUH, W. BAI, W. SHI, G. PENNEY et D. RUECKERT : *Probabilistic Edge Map (PEM) for 3D Ultrasound Image Registration and Multi-atlas Left Ventricle Segmentation*, pages 223–230. Springer International Publishing, Cham, 2015. ISBN 978-3-319-20309-6.
- [185] F. ORDERUD et S.I RABBEN : Real-time 3D segmentation of the left ventricle using deformable subdivision surfaces. *In Computer Vision and Pattern Recognition (CVPR)*, 2008.
- [186] N. F. OSMAN, W. S. KERWIN, E. R. MCVEIGH et J. L. PRINCE : Cardiac motion tracking using cine harmonic phase (harp) magnetic resonance imaging. *Magn Reson Med*, 42(6):1048–1060, Dec 1999.
- [187] N. F. OSMAN, E. MCVEIGH et J. PRINCE : Imaging heart motion using harmonic phase mri. *IEEE transactions on medical imaging*, 19:186–202, 04 2000.

-
- [188] J. D. O’SULLIVAN : A fast sinc function gridding algorithm for fourier inversion in computer tomography. *IEEE Transactions on Medical Imaging*, 4(4):200–207, 1985.
- [189] L. PAN, J. L. PRINCE, J. A. C. LIMA et N. F. OSMAN : Fast tracking of cardiac motion using 3d-harp. *IEEE Transactions on Biomedical Engineering*, 52(8):1425–1435, 2005.
- [190] A. PAPACHRISTIDIS, M. GELEIJNSE, E. GALLI, B. HEYDE, M. ALESSANDRINI, D. BARBOSA, E. DONAL, M. MONAGHAN, O. BERNARD, J. D’HOOGHE et J. G. BOSCH : Clinical Expert Delineation of 3D Left Ventricular Echocardiograms for the CETUS Segmentation Challenge. In *Proceedings MICCAI Challenge on Echocardiographic Three-Dimensional Ultrasound Segmentation (CETUS)*, Boston, Sept. 14, pages 9–16, 2014.
- [191] C. PAPADACCI, M. PERNOT, M. COUADE, M. FINK et M. TANTER : High-contrast ultrafast imaging of the heart. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 61(2):288–301, 2014.
- [192] X. PAPADEMETRIS, A. SINUSAS, D. P. DIONE, R. CONSTABLE et J. S. DUNCAN : Estimation of 3d left ventricular deformation from medical images using biomechanical models. *IEEE transactions on medical imaging*, 21:786–800, 08 2002.
- [193] A. PASQUET, A. C. POULEUR, D. VANCRAEYNES, J. L. VANOVERSHELDE, M. S. AMZULESCU, B. L. GERBER, M. De CRAENE et H. LANGET : Myocardial strain imaging : review of general principles, validation, and sources of discrepancies. *European Heart Journal*, 03 2019.
- [194] J. PATRAVALI, S. JAIN et S. CHILAMKURTHY : 2d-3d fully convolutional neural networks for cardiac mr segmentation. In *Proc. STACOM-MICCAI, LNCS, volume 10663*, pages 130–139. 2017.
- [195] G. PEDRIZZETTI, P. CLAUS, P. J. KILNER et E. NAGEL : Principles of cardiovascular magnetic resonance feature tracking and echocardiographic speckle tracking for informed clinical use. *Journal of Cardiovascular Magnetic Resonance*, 18(1):51, Aug 2016. ISSN 1532-429X. URL <https://doi.org/10.1186/s12968-016-0269-7>.
- [196] J. PEDROSA, D. BARBOSA, B. HEYDE, F. SCHNELL, A. RÖSNER, P. CLAUS et J. D’HOOGHE : Left Ventricular Myocardial Segmentation in 3-D Ultrasound Recordings : Effect of Different Endocardial and Epicardial Coupling Strategies. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 64(3):525–536, March 2017.
- [197] J. PEDROSA, S. QUEIRÓS, O. BERNARD, J. ENGVALL, T. EDVARSEN, E. NAGEL et J. D’HOOGHE : Fast and Fully Automatic Left Ventricular Segmentation and Tracking in Echocardiography Using Shape-Based B-Spline Explicit Active Surfaces. *IEEE Transactions on Medical Imaging*, 36(11):2287–2296, 2017.
- [198] D. PERPERIDIS, R. H. MOHIADDIN et D. RUECKERT : Spatio-temporal free-form registration of cardiac mr image sequences. *Medical image analysis*, 9:441–56, 11 2005.
- [199] S. PETERSEN, P. M. MATTHEWS, J. M. FRANCIS, M. D. ROBSON, F. ZEMRAK, R. BOUBERTAKH, A. A. YOUNG *et al.* : Uk biobank’s cardiovascular magnetic resonance protocol. *Journal of Cardiovascular Magnetic Resonance*, 18(8), 2016.
- [200] C. PETITJEAN et J. DACHER : A review of segmentation methods in short axis cardiac MR images. *Med. Image Anal.*, 15(2):169–184, 2011. ISSN 1361-8415.

- [201] C. PETITJEAN, M. A. ZULUAGA, W. BAI, J. N. DACHER, D. GROSGEORGE, J. CAUDRON, S. RUAN, I. B. AYED, M. J. CARDOSO *et al.* : Right ventricle segmentation from cardiac MRI : A collation study. *Medical Image Analysis*, 19(1):187–202, 2015.
- [202] K.K. POPPE *et al.* : Ethnic-specific normative reference values for echocardiographic la and lv size, lv mass, and systolic function : The echonormal study. *JACC : Cardiovascular Imaging*, 8(6):656 – 665, 2015. ISSN 1936-878X. URL <http://www.sciencedirect.com/science/article/pii/S1936878X15002193>.
- [203] J. POREE, D. POSADA, A. HODZIC, F. TOURNOUX, G. CLOUTIER et D. GARCIA : High-frame-rate echocardiography using coherent compounding with doppler-based motion-compensation. *IEEE Transactions on Medical Imaging*, PP(99):1–1, 2016.
- [204] R. P. POUDEL, P. LAMATA et G. MONTANA : Recurrent fully convolutional neural networks for multi-slice MRI cardiac segmentation. *arXiv :1608.03974*, 2016.
- [205] A. PRAKOSA, M. SERMESANT, H. DELINGETTE, S. MARCHESSEAU, E. SALOUX, P. ALLAIN, N. VILLAIN et N. AYACHE : Generation of synthetic but visually realistic time series of cardiac images combining a biophysical model and clinical images. *Medical Imaging, IEEE Transactions on*, 32(1):99–109, 2013.
- [206] A. PRAKOSA, M. SERMESANT, H. DELINGETTE, S. MARCHESSEAU, E. SALOUX, P. ALLAIN, N. VILLAIN et N. AYACHE : Generation of synthetic but visually realistic time series of cardiac images combining a biophysical model and clinical images. *IEEE Transactions on Medical Imaging*, 32(1):99–109, Jan 2013. ISSN 0278-0062.
- [207] F. PRIEUR, B. DÉNARIÉ, A. AUSTENG et H. TORP : Correspondence - multi-line transmission in medical imaging using the second-harmonic signal. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 60(12):2682–2692, Dec 2013.
- [208] J. PROVOST, V.T.H NGUYEN, D. LEGRAND, S. OKRASINSKI, A. COSTET, A. GAMBHIR, H. GARAN et E.E. KONOFAGOU : Electromechanical wave imaging for arrhythmias. *Physics in medicine and biology*, 56(22):L1, 2011.
- [209] J. PROVOST, C. PAPADACCI, J. E. ARANGO, M. IMBAULT, M. FINK, J. L. GENNISSON, M. TANTER et M. PERNOT : 3d ultrafast ultrasound imaging in vivo. *Physics in medicine and biology*, 59(19):L1, 2014.
- [210] J. PROVOST, C. PAPADACCI, C. DEMENE, J. L. GENNISSON, M. TANTER et M. PERNOT : 3-d ultrafast doppler imaging applied to the noninvasive mapping of blood vessels in vivo. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 62:1467–1472, 2015.
- [211] Z. QIAN, Q. LIU, D. N. METAXAS et L. AXEL : Identifying regional cardiac abnormalities from myocardial strains using nontracking-based strain estimation and spatio-temporal tensor analysis. *IEEE Transactions on Medical Imaging*, 30(12):2017–2029, 2011.
- [212] S. QUEIRÓS, A. PAPACHRISTIDIS, D. BARBOSA, K. C. THEODOROPOULOS, J. C. FONSECA, M. J. MONAGHAN, J. L. VILAÇA et J. D’HOOGE : Aortic valve tract segmentation from 3D-TEE using shape-based B-spline explicit active surfaces. *IEEE Transactions on Medical Imaging*, 35(9):2015–2025, Sept 2016. ISSN 0278-0062.

-
- [213] S. QUEIRÓS, D. BARBOSA, B. HEYDE, P. MORAIS, J. L. VILAÇA, D. FRIBOULET, O. BERNARD et J. D’HOOGHE : Fast automatic myocardial segmentation in 4D cine CMR datasets. *Medical Image Analysis*, 18(7):1115 – 1131, 2014.
- [214] S. QUEIRÓS, J. L. VILAÇA, P. MORAIS, J. C. FONSECA, J. D’HOOGHE et D. BARBOSA : Fast left ventricle tracking using localized anatomical affine optical flow. *International Journal for Numerical Methods in Biomedical Engineering*, 33(11):e2871. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/cnm.2871>. e2871 cm.2871.
- [215] P. RADAU, Y. LU, K. CONNELLY, G. PAUL, A. DICK et G. WRIGHT : Evaluation framework for algorithms segmenting short axis cardiac MRI. In *The MIDAS Journal - Cardiac MR Left Ventricle Segmentation Challenge*, 2009.
- [216] F. RADEMAKERS, J. ENGVALL, T. EDVARDSEN, M. MONAGHAN, R. SICARI, E. NAGEL, J. ZAMORANO, H. UKKONEN, T. EBBERS, V. DI BELLO, J. VOIGT, L. HERBOTS, P. CLAUS et J. D’HOOGHE : Determining optimal noninvasive parameters for the prediction of left ventricular remodeling in chronic ischemic patients. *Scandinavian Cardiovascular Journal*, 47(6):329–334, décembre 2013. ISSN 1651-2006.
- [217] K. RAJPOOT, V. GRAU et J. A. NOBLE : Local-phase based 3d boundary detection using monogenic signal and its application to real-time 3-d echocardiography images. In *2009 IEEE International Symposium on Biomedical Imaging : From Nano to Macro*, pages 783–786, June 2009.
- [218] K. RAJPOOT, V. GRAU, J. A. NOBLE, H. BECHER et C. SZMIGIELSKI : The evaluation of single-view and multi-view fusion 3D echocardiography using image-driven segmentation and tracking. *Medical Image Analysis*, 15(4):514 – 528, 2011.
- [219] S. A. REISNER, P. LYSYANSKY, Y. AGMON, D. MUTLAK, J. LESSICK et Z. FRIEDMAN : Global longitudinal strain : A novel index of left ventricular systolic function. *Journal of the American Society of Echocardiography : official publication of the American Society of Echocardiography*, 17:630–3, 06 2004.
- [220] U. REITER, G. REITER, K. DORR, A. GREISER, R. MADERTHANER et M. FUCHSJÄGER : Normal diastolic and systolic myocardial t1 values at 1.5-t mr imaging : correlations and blood normalization. *Radiology*, 271(2):365–372, 2013.
- [221] S. RICCI, L. BASSI et P. TORTOLI : Real-time vector velocity assessment through multigate doppler and plane waves. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 61(2):314–324, 2014.
- [222] S. V. RICHARD et J. A. NOBLE : Delineating anatomical boundaries using the boundary fragment model. *Medical Image Analysis*, 17(8):1123–1136, 2013.
- [223] M-M. ROHE, M. SERMESANT et X. PENNEC : Automatic multi-atlas segmentation of myocardium with svf-net. In *Proc. STACOM-MICCAI, LNCS, volume 10663*, pages 170–177, 2017.
- [224] M.M ROHE, M. SERMESANT et X. PENNEC : Svf-net : Learning deformable image registration using shape matching. In *Proc. MICCAI*, 2017.
- [225] O. RONNEBERGER, P. FISCHER et T. BROX : U-Net : Convolutional Networks for Biomedical Image Segmentation. In *Proc. MICCAI*, pages 234–241. 2016.

- [226] D ROPERS : Detection of coronary artery stenoses with thin-slice multi-detector row spiral computed tomography and multiplanar reconstruction. *Circulation*, 107:664–666, 02 2003.
- [227] G. A. ROTH, M. H. FOROUZANFAR, A. E. MORAN, R. BARBER, G. NGUYEN, V. L. FEIGIN, M. NAGHAVI, G. A. MENSAH et C. J. L. MURRAY : Demographic and epidemiologic drivers of global cardiovascular mortality. *New England Journal of Medicine*, 372(14):1333–1341, 2015.
- [228] C. RUPPRECHT, E. HUAROC, M. BAUST et N. NAVAB : Deep active contours. *arXiv :1607.05074*, 2016.
- [229] A. RUTZ, S. RYF, S. PLEIN, P. BOESIGER et S. KOZERKE : Accelerated whole-heart 3d cspamm for myocardial motion quantification. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, 59:755–63, 04 2008.
- [230] A. K. RUTZ, S. RYF, S. PLEIN, P. BOESIGER et S. KOZERKE : Accelerated whole-heart 3d cspamm for myocardial motion quantification. *Magnetic Resonance in Medicine*, 59 (4):755–763, 2008.
- [231] J. SAINTE-MARIE, D. CHAPELLE, R. CIMRMAN et M. SORINE : Modeling and estimation of the cardiac electromechanical activity. *Computers & Structures*, 84(28):1743 – 1759, 2006. ISSN 0045-7949. URL <http://www.sciencedirect.com/science/article/pii/S0045794906001878>.
- [232] H SAKUMA, N FUJITA, T K FOO, G R CAPUTO, S J NELSON, J HARTIALA, A SHIMAKAWA et C B HIGGINS : Evaluation of left ventricular volume and mass with breath-hold cine mr imaging. *Radiology*, 188(2):377–380, 1993.
- [233] S. SALLES, A.J.Y. CHEE, D. GARCIA, A.C.H. YU, D. VRAY et H. LIEBGOTT : 2-d arterial wall motion imaging using ultrafast ultrasound and transverse oscillations. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 62(6):1047–1058, 2015.
- [234] J. SANDER, B. D. de VOS, J. M. WOLTERINK et I. IŞGUM : Towards increased trustworthiness of deep learning segmentation methods on cardiac mri. In *SPIE Medical Imaging*, 2019.
- [235] K. SCHEFFLER et S. LEHNHARDT : Principles and applications of balanced ssfp techniques. *European radiology*, 13(11):2409–2418, 2003.
- [236] Klaus SCHEFFLER et Stefan LEHNHARDT : Principles and applications of balanced ssfp techniques. *European radiology*, 13(11):2409–2418, 2003.
- [237] WP SEGARS, G STURGEON, S MENDONCA, Jason GRIMES et Benjamin MW TSUI : 4d xcat phantom for multimodality imaging research. *Medical physics*, 37(9):4902–4915, 2010.
- [238] M. SERMESANT, P. MOIREAU, O. CAMARA, J. SAINTE-MARIE, R. ANDRIANTSIMIYAVONA, R. CIMRMAN, D. L. G. HILL, D. CHAPELLE et R. RAZAVI : Cardiac function estimation from mri using a heart model and data assimilation : advances and difficulties. *Medical Image Analysis*, 10(4):642–656, 2006.

-
- [239] D. P. SHATTUCK, M. D. WEINSHENKER, S. W. SMITH et O. T. von RAMM : Explososcan : A parallel processing technique for high speed ultrasound imaging with linear phased arrays. *The Journal of the Acoustical Society of America*, 75(4):1273–1282, 1984.
- [240] P. SHI, A. SINUSAS, R. CONSTABLE et J. S. DUNCAN : Volumetric deformation analysis using mechanics-based data fusion : Applications in cardiac motion recovery. *International Journal of Computer Vision*, 35:87–107, 11 1999.
- [241] K. SHIINO, A. YAMADA, M. ISCHENKO, B. K. KHANDHERIA, M. HUDAVERDI, V. SPERANZA, M. HARTEN, A. BENJAMIN, C. R. HAMILTON-CRAIG, D. G. PLATTS, D. J. BURSTOW, G. M. SCALIA et J. CHAN : Intervendor consistency and reproducibility of left ventricular 2d global and regional strain with two different high-end ultrasound systems. *European Heart Journal - Cardiovascular Imaging*, 18(6):707–716, 2017. URL <http://dx.doi.org/10.1093/ehjci/jew120>.
- [242] G. K. SINGH, B. CUPPS, M. PASQUE, P. K. WOODARD, M. R. HOLLAND et A. LUDOMIRSKY : Accuracy and reproducibility of strain by speckle tracking in pediatric subjects with normal heart and single ventricular physiology : A two-dimensional speckle-tracking echocardiography and magnetic resonance imaging correlative study. *Journal of the American Society of Echocardiography*, 23(11):1143–1152, Nov 2010. ISSN 0894-7317. URL <https://doi.org/10.1016/j.echo.2010.08.010>.
- [243] O. A. SMISETH, H. TORP, A. OPDAHL, K. H. HAUGAA et S. URHEIM : Myocardial strain imaging : how useful is it in clinical decision making? *European Heart Journal*, 37(15):1196–1207, 2016. URL <http://dx.doi.org/10.1093/eurheartj/ehv529>.
- [244] E. SMISTAD et F. LINDSETH : Real-time tracking of the left ventricle in 3d ultrasound using kalman filter and mean value coordinates. In *Proc. MICCAI Challenge on Echocardiographic Three-Dimensional Ultrasound Segmentation (CETUS)*, Boston, MIDAS Journal, pages 65–72. 2014.
- [245] E. SMISTAD, A. ØSTVIK, I. Mjaland SALTE, S. LECLERC, O. BERNARD et L. LOVSTAKKEN : Fully automatic real-time ejection fraction and MAPSE measurements in 2D echocardiography using deep neural networks. In *2018 IEEE International Ultrasonics Symposium (IUS)*, pages 1–1, Oct 2018.
- [246] E. SMISTAD, A. ØSTVIK, B. O. HAUGEN et L. LOVSTAKKEN : 2D left ventricle segmentation using deep learning. In *2017 IEEE International Ultrasonics Symposium (IUS)*, pages 1–4, 2017.
- [247] O. SOMPHONE, M. De CRAENE, R. ARDON, B. MORY, P. ALLAIN, H. GAO, J. D’HOOGHE, S. MARCHESSEAU, M. SERMESANT, H. DELINGETTE *et al.* : Fast myocardial motion and strain estimation in 3d cardiac ultrasound with sparse demons. In *Biomedical Imaging (ISBI), 2013 IEEE 10th International Symposium on*, pages 1182–1185. IEEE, 2013.
- [248] H. STARK, J. W. WOODS, I. PAUL et R. HINGORANI : An investigation of computerized tomography by direct fourier inversion and optimum interpolation. *IEEE Transactions on Biomedical Engineering*, BME-28(7):496–505, 1981.
- [249] M. G. STRINTZIS et I. KOKKINIDIS : Maximum likelihood motion estimation in ultrasound image sequences. *IEEE Signal Processing Letters*, 4(6):156–157, June 1997. ISSN 1070-9908.

- [250] M. SUHLING, M. ARIGOVINDAN, C. JANSEN, P. HUNZIKER et M. UNSER : Myocardial motion analysis from b-mode echocardiograms. *IEEE Transactions on Image Processing*, 14(4):525–536, April 2005. ISSN 1057-7149.
- [251] A. SUINESIAPUTRA, B. R. COWAN, A. O. AL-AGAMY, M. A. ELATTAR, N. AYACHE, A. S. FAHMY, A. M. KHALIFA, P. MEDRANO-GRACIA, M. JOLLY, A. H. KADISH, D. C. LEE, J. MARGETA, S. K. WARFIELD et A. A. YOUNG : A collaborative resource to build consensus for automated left ventricular segmentation of cardiac mr images. *Med. Image Anal.*, 18(1):50–62, 2014.
- [252] H. SUNDAR, H. LITT et D. SHEN : Estimating myocardial motion by 4d image warping. *Pattern recognition*, 42:2514–2526, 11 2009.
- [253] L. K. TAN, Y. M. LIEW, E. LIM et R. A. MCCLAUGHLIN : Cardiac left ventricle segmentation using convolutional neural network regression. *In Proc. IECBES*, pages 490–93, 2016.
- [254] M. TANTER et M. FINK : Ultrafast imaging in biomedical ultrasound. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 61(1):102–119, January 2014.
- [255] V. TAVAKOLI et N. SAHBA : Cardiac motion and strain detection using 4d ct images : comparison with tagged mri, and echocardiography. *The International Journal of Cardiovascular Imaging*, 30(1):175–184, Jan 2014.
- [256] R. J. TAYLOR, W. E. MOODY, F. UMAR, N. C. EDWARDS, T. J. TAYLOR, B. STEGEMANN, J. N. TOWNEND, K. N. HOR, R. P. STEEDS, W. MAZUR et F. LEYVA : Myocardial strain measurement with feature-tracking cardiovascular magnetic resonance : normal values. *European Heart Journal - Cardiovascular Imaging*, 16(8):871–881, 2015.
- [257] J.-P. THIRION : Image matching as a diffusion process : an analogy with maxwell’s demons. *Medical Image Analysis*, 2(3):243 – 260, 1998.
- [258] C. TOBON-GOMEZ, M. De CRAENE, K. MCLEOD, L. TAUTZ, W. SHI, A. HENNEMUTH, A. PRAKOSA, H. WANG, G. CARR-WHITE, S. KAPETANAKIS *et al.* : Benchmarking framework for myocardial tracking and deformation algorithms : An open access database. *Medical Image Analysis*, 17(6):632–648, 2013.
- [259] C TOBON-GOMEZ, FM SUKNO, BH BIJNENS, M HUGUET et AF FRANGI : Realistic simulation of cardiac magnetic resonance studies modeling anatomical variability, trabeculae, and papillary muscles. *Magnetic Resonance in Medicine*, 65(1):280–288, 2011.
- [260] L. TONG, H. GAO et J. D’HOOGHE : Multi-transmit beam forming for fast cardiac imaging—a simulation study. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 60(8):1719–1731, August 2013.
- [261] L. TONG, A. RAMALLI, R. JASAITYTE, P. TORTOLI et J. D’HOOGHE : Multi-transmit beam forming for fast cardiac imaging—experimental validation and in vivo application. *IEEE Transactions on Medical Imaging*, 33(6):1205–1219, June 2014.
- [262] B. TOUIL, A. BASARAB, P. DELACHARTRE, O. BERNARD et D. FRIBOULET : Analysis of motion tracking in echocardiographic image sequences : Influence of system geometry and point-spread function. *Ultrasonics*, 50(3):373 – 386, 2010. ISSN 0041-624X. URL <http://www.sciencedirect.com/science/article/pii/S0041624X09001024>.

-
- [263] G. E. TRAHEY et L. F. NOCK : Synthetic receive aperture imaging with phase correction for motion and for tissue inhomogeneities. ii. effects of and correction for motion. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, 39(4):496–501, 1992.
- [264] D. TRAN, L. BOURDEV, R. FERGUS, L. TORRESANI et M. PALURI : Learning spatiotemporal features with 3d convolutional networks. *In Proc. ICCV*, 2015.
- [265] P. V. TRAN : A fully convolutional neural network for cardiac segmentation in short-axis MRI. *arXiv :1604.00494*, 2017.
- [266] Y. TSADOK, Z. FRIEDMAN, B. A. HALUSKA, R. HOFFMANN et D. ADAM : Myocardial strain assessment by cine cardiac magnetic resonance imaging using non-rigid registration. *Magnetic Resonance Imaging*, 34, 12 2015.
- [267] G TZIRITAS et E GRINIAS : Fast fully-automatic localization of left ventricle and myocardium in mri using mrf model optimization, substructures tracking and b-spline smoothing. *In Proc. STACOM-MICCAI, LNCS, volume 10663*, pages 91–100, 2017.
- [268] J. UDESEN, F. GRAN, K. L. HANSEN, J. A. JENSEN, C. THOMSEN et M. B. NIELSEN : High frame-rate blood vector velocity imaging using plane waves : simulations and preliminary experiment. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 55 (8):1729–1743, 2008.
- [269] J. ULEN, P. STRANDMARK et F. KAHL : An efficient optimization framework for multi-region segmentation based on lagrangian duality. *IEEE Trans. Med. Imaging*, 32(2):178–188, 2013. ISSN 0278-0062.
- [270] M. UZUMCU, R. van der GEEST, C. SWINGEN, J. REIBER et B. LELIEVELDT : Time continuous tracking and segmentation of cardiovascular magnetic resonance images using multidimensional dynamic programming. *Invest Radiol*, 411:52–62, 2006.
- [271] M. van STRALEN, A. HAAK, K. LEUNG, G. van BURKEN et J. BOSCH : Segmentation of multi-center 3d left ventricular echocardiograms by active appearance models. *In Proc. MICCAI Challenge on Echocardiographic Three-Dimensional Ultrasound Segmentation (CETUS), Boston, MIDAS Journal*, pages 73–80. 2014.
- [272] M. van STRALEN, A. HAAK, K. Y. E. LEUNG, G. van BURKEN, C. BOS et J. G. BOSCH : Full-cycle left ventricular segmentation and tracking in 3D echocardiography using active appearance models. *In 2015 IEEE International Ultrasonics Symposium (IUS)*, pages 1–4, Oct 2015.
- [273] M. van STRALEN, K. Y. E. LEUNG, M. M. VOORMOLEN, N. de JONG, A. F. W. van der STEEN, J. H. C. REIBER et J. G. BOSCH : Automatic segmentation of the left ventricle in 3d echocardiography using active appearance models. *In IEEE Ultrasonics Symposium*, pages 1480–1483, 2007.
- [274] M. van STRALEN, K.Y.E. LEUNG, M.M. VOORMOLEN, N. de JONG, A.F.W. van der STEEN, J.H.C. REIBER et J.G. BOSCH : Time continuous detection of the left ventricular long axis and the mitral valve plane in 3-D echocardiography. *Ultrasound in Medicine and Biology*, 34(2):196 – 207, 2008. ISSN 0301-5629.

- [275] O. VILLEMMAIN, M. CORREIA, D. KHRAICHE, I. PODETTI, M. MEOT, A. LEGENDRE, M. TANTER, D. BONNET et M. PERNOT : Myocardial stiffness assessment using shear wave imaging in pediatric hypertrophic cardiomyopathy. *JACC : Cardiovascular Imaging*, 11(5):779 – 781, 2018.
- [276] O. T. VON RAMM, S. W. SMITH et H. G. PAVY : High-speed ultrasound volumetric imaging system. ii. parallel processing and image display. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 38(2):109–115, March 1991.
- [277] R. F. WAGNER, S. W. SMITH, J. M. SANDRIK et H. LOPEZ : Statistics of speckle in ultrasound b-scans. *IEEE Transactions on Sonics and Ultrasonics*, 30(3):156–163, May 1983. ISSN 0018-9537.
- [278] E. WAKS, J. L. PRINCE et A. S. DOUGLAS : Cardiac motion simulator for tagged mri. *In Mathematical Methods in Biomedical Image Analysis, 1996., Proceedings of the Workshop on*, pages 182–191. IEEE, 1996.
- [279] H. WANG et A. A. AMINI : Cardiac motion and deformation recovery from mri : A review. *IEEE Transactions on Medical Imaging*, 31(2):487–503, Feb 2012.
- [280] H. WANG, M. KADBI, M. KOTYS, M. ERSOY, G. P. CHATZIMAVROUDIS, R. M. SETSER, M. ALSHAHER, S. E. FISCHER et A. A. AMINI : Orthogonal cspamm (ocspamm) mr tagging for imaging ventricular wall motion. *In 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 535–538. IEEE, 2011.
- [281] J. WANG et J. LU : Motion artifacts of extended high frame rate imaging. *Ultrasonics, Ferroelectrics, and Frequency Control, IEEE Transactions on*, 54(7):1303–1315, 2007.
- [282] S. K. WARFIELD, K. H. ZOU et W. M. WELLS : Simultaneous truth and performance level estimation (staple) : an algorithm for the validation of image segmentation. *IEEE Trans. Med. Imaging*, 23(7):903–921, July 2004. ISSN 0278-0062.
- [283] H. D. WHITE, R. M. NORRIS, M. A. BROWN, P. W. BRANDT, R. M. WHITLOCK et C. J. WILD : Left ventricular end-systolic volume as the major determinant of survival after recovery from myocardial infarction. *Circulation.*, 76:44–51, 1987.
- [284] M.C. Van WIJK et J.M. THIJSSSEN : Performance testing of medical ultrasound equipment : Fundamental vs. harmonic mode. *Ultrasonics*, 40:585—591, 2002.
- [285] L. WISSMANN, C. SANTELLI, W. P. SEGARS et S. KOZERKE : Mrxcats : Realistic numerical phantoms for cardiovascular magnetic resonance. *Journal of Cardiovascular Magnetic Resonance*, 16(1):1, 2014.
- [286] J. M WOLTERINK, T. LEINER, M. A VIERGEVER et I. ISGUM : Automatic segmentation and disease classification using cardiac cine mr images. *In Proc. STACOM-MICCAI, LNCS, volume 10663*, pages 101–110, 2017.
- [287] H. YANG, J. SUN, H. LI, L. WANG et Z. XU : Deep fusion net for multi-atlas segmentation : Application to cardiac MR images. *In Proc. MICCAI*, pages 521–528. 2016.
- [288] L. YANG, B. GEORGESCU, Y. ZHENG, Y. WANG, P. MEER et D. COMANICIU : Prediction based collaborative trackers (PCT) : A robust and accurate approach toward 3D medical

-
- object tracking. *IEEE Transactions on Medical Imaging*, 30(11):1921–1932, 2011. ISSN 0278-0062.
- [289] X. YANG, C. BIAN, L. YU, D. NI et P-A. HENG : Class-balanced deep neural network for automatic ventricular structure segmentation. *In Proc. STACOM-MICCAI, LNCS, volume 10663*, pages 152–160, 2017.
- [290] B. Y. S. YIU, I. K. H. TSANG et C. H. ALFRED : A modified synthetic aperture imaging approach with axial motion compensation. *In 2008 IEEE Ultrasonics Symposium*, pages 1254–1257. IEEE, 2008.
- [291] A. YOUNG : Model tags : Direct three-dimensional tracking of heart wall motion from tagged magnetic resonance images. *Medical image analysis*, 3:361–72, 2000.
- [292] A. A. YOUNG, D. L. KRAITCHMAN, L. DOUGHERTY et L. AXEL : Tracking and finite element analysis of stripe deformation in magnetic resonance tagging. *IEEE Transactions on Medical Imaging*, 14(3):413–421, 1995.
- [293] E. A. ZERHOUNI, D. M. PARISH, W. J. ROGERS, A. YANG et E. P. SHAPIRO : Human heart : tagging with mr imaging—a method for noninvasive assessment of myocardial motion. *Radiology*, 169(1):59–63, 1988.
- [294] L. ZHANG, A. GOOYA, B. DONG, R. HUA, S. E. PETERSEN, P. MEDRANO-GRACIA et A. F. FRANGI : Automated quality assessment of cardiac mr images using convolutional neural networks. *In Proc. SASHIMI-MICCAI*, pages 138–145. 2016.
- [295] M. ZHANG, A. BESSON, R. E. CARRILLO, F. VARRAY, H. LIEBGOTT, J.P. THIRAN, D. FRIBOULET et O. BERNARD : Extension of ultrasound fourier slice imaging theory to sectorial acquisition. *In IEEE International Ultrasonics Symposium (IUS)*, 2015.
- [296] M. ZHANG, H. LIEBGOTT, F. VARRAY, D. FRIBOULET et O. BERNARD : A Fourier-based formalism for 3D ultrafast imaging with diverging waves. *In 2016 IEEE International Ultrasonics Symposium (IUS)*, pages 1–4, Sept 2016.
- [297] M. ZHANG, H. LIEBGOTT, F. VARRAY, D. FRIBOULET et O. BERNARD : Fourier-based ultrafast ultrasound imaging based on in-phase quadrature (IQ) data. *In 2017 IEEE International Ultrasonics Symposium (IUS)*, pages 1–1, Sept 2017.
- [298] M. ZHANG, F. VARRAY, L. PETRUSCA, D. FRIBOULET, H. LIEBGOTT et O. BERNARD : Fourier-based 3D ultrafast ultrasound imaging with diverging waves : In vitro experiment validation. *In 2017 IEEE International Ultrasonics Symposium (IUS)*, pages 1–1, Sept 2017.
- [299] Y. ZHOU, O. BERNARD, E. SALOUX, A. MANRIQUE, P. ALLAIN, S. MAKRAM-EBEID et M. De CRAENE : 3d harmonic phase tracking with anatomical regularization. *Medical Image Analysis*, 26(1):70–81, 2015.
- [300] Y. ZHOU, M. De CRAENE, O. SOMPHONE, M. SERMESANT et O. BERNARD : Generation of realistic 4d synthetic cspamm tagged mr sequences for benchmarking cardiac motion tracking algorithms. *In International Workshop on Simulation and Synthesis in Medical Imaging*, pages 212–219. Springer, in press.

- [301] Y. ZHOU, S. GIFFARD-ROISIN, M. De CRAENE, S. CAMARASU-POP, J. D'HOOGHE, M. ALESSANDRINI, D. FRIBOULET, M. SERMESANT et O. BERNARD : A Framework for the Generation of Realistic Synthetic Cardiac Ultrasound and Magnetic Resonance Imaging Sequences From the Same Virtual Patients. *IEEE Transactions on Medical Imaging*, 37 (3):741–754, March 2018.
- [302] Z. ZHOU, M.R. SIDDIQUEE, N. TAJBAKHSI et J. LIANG : Unet++ : A nested u-net architecture for medical image segmentation. *In in proc. of Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pages 3–11. 2018.
- [303] X. ZHUANG, C. YAO, Y. MA, D. HAWKES, G. PENNEY et S. OURSELIN : Registration-based propagation for whole heart segmentation from compounded 3D echocardiography. *In IEEE International Symposium on Biomedical Imaging*, pages 1093–1096, 2010.
- [304] C. ZOTTI, Z. LUO, O. HUMBERT, A. LALANDE et P. M. JODOIN : Gridnet with automatic shape prior registration for automatic mri cardiac segmentation. *In Proc. STACOM-MICCAI, LNCS, volume 10663*, pages 73–81, 2017.