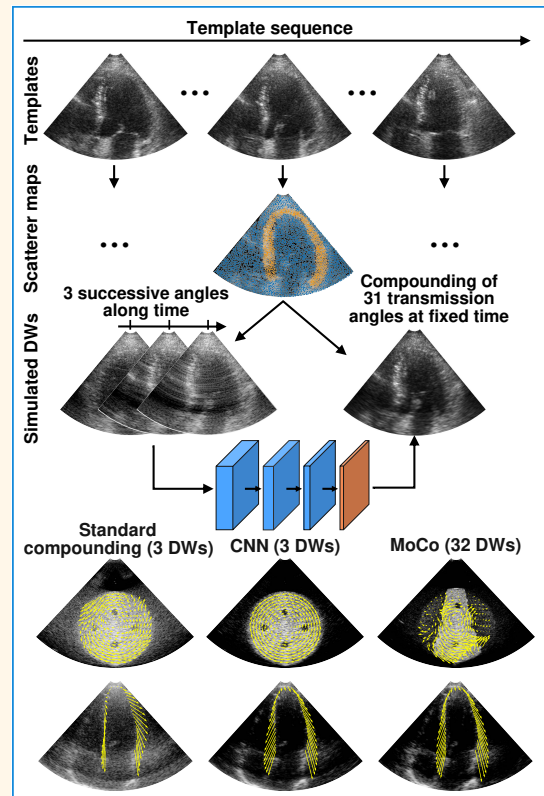


# Ultrafast Cardiac Imaging Using Deep Learning For Speckle-Tracking Echocardiography

Jingfeng Lu, Fabien Millioz, François Varray, Jonathan Porée, Jean Provost, Olivier Bernard, Damien Garcia, and Denis Friboulet

**Abstract**— High-quality ultrafast ultrasound imaging is based on coherent compounding from multiple transmissions of plane waves (PW) or diverging waves (DW). However, compounding results in reduced frame rate, as well as destructive interferences from high-velocity tissue motion if motion compensation (MoCo) is not considered. While many studies have recently shown the interest of deep learning for the reconstruction of high-quality static images from PW or DW, its ability to achieve such performance while maintaining the capability of tracking cardiac motion has yet to be assessed. In this paper, we addressed such issue by deploying a complex-weighted convolutional neural network (CNN) for image reconstruction and a state-of-the-art speckle tracking method. The evaluation of this approach was first performed by designing an adapted simulation framework, which provides specific reference data, i.e. high quality, motion artifact-free cardiac images. The obtained results showed that, while using only three DWs as input, the CNN-based approach yielded an image quality and a motion accuracy equivalent to those obtained by compounding 31 DWs free of motion artifacts. The performance was then further evaluated on non-simulated, experimental *in vitro* data, using a spinning disk phantom. This experiment demonstrated that our approach yielded high-quality image reconstruction and motion estimation, under a large range of velocities and outperforms a state-of-the-art MoCo-based approach at high velocities. Our method was finally assessed on *in vivo* datasets and showed consistent improvement in image quality and motion estimation compared to standard compounding. This demonstrates the feasibility and effectiveness of deep learning reconstruction for ultrafast speckle-tracking echocardiography.

**Index Terms**— Echocardiography, ultrafast ultrasound imaging, diverging wave, deep learning, motion estimation, speckle tracking



This work was supported in part by Sichuan University Postdoctoral Interdisciplinary Innovation Fund under Grant JCXK2232, and partially supported by the framework of the LABEX PRIMES (ANR-11-LABX-0063) of Université de Lyon and the LABEX CELYA (ANR-11-LABX-0060) of Université de Lyon, both within the program "Investissements d'Avenir" (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR). The *in vivo* part was partially supported by a Tremplin International - INSERM grant (DELEA-2022-480/DELCP-2022-271).

J. Lu is affiliated with the School of Cyber Science and Engineering, Sichuan University, Chengdu, China, he was with the School of Instrumentation Science and Engineering, Harbin Institute of Technology, Harbin, China at the time of writing this paper.

F. Millioz, F. Varray, O. Bernard, D. Garcia, and D. Friboulet are affiliated with the University of Lyon, CREATIS, CNRS UMR 5220, Inserm U1294, INSA-Lyon, University of Lyon 1, Villeurbanne, France.

J. Porée and J. Provost are affiliated with Engineering Physics Department, Polytechnique Montréal, Montréal, Canada.

The proposed model with trained weights in Open Neural Network Exchange (ONNX) format and a demo script are available at <https://github.com/Jingfeng-LU/Deep-Ultrafast-Echocardiography>

## I. INTRODUCTION

CONVENTIONAL echocardiography typically reaches 30 to 100 frames/second with a line-by-line ultrasound system [1], where multiple sectors of the entire field of view are insonified using sequential narrow beams. Such a frame rate is sufficient for speckle tracking echocardiography in a resting adult whose heartbeat is approximately 70 per minute [2]. However, higher frame rates may be required in some imaging situations, e.g. i) stress echocardiography [3], ii) all-four-chamber strain imaging [4], [5], iii) electromechanical wave imaging [6], and iv) myocardial shear wave imaging [7], [8].

To achieve higher frame rates, wide beam imaging systems using plane wave (PW) [9] or diverging wave (DW) [10]

### Highlights

- **A novel framework for deep learning-based ultrafast ultrasound cardiac imaging, and evaluation of its temporal coherence for speckle-tracking echocardiography.**
- **The proposed approach demonstrates high image quality while preserving consistent speckle pattern for tracking high-speed tissue motion, when applied on simulated, *in-vitro*, and *in-vivo* data.**
- **This approach has the potential to improve cardiac diagnosis using dynamic analysis at ultra-high frame rates (1500 frames/second in our experiments), owing to its high imaging quality and detectable tissue speed.**

have been developed to reconstruct the full field of view in a single transmit. Such schemes can provide frame rates up to several thousands of frames/second. Ultrafast imaging can thus capture short-duration events in cardiac cycles, potentially improving the accuracy of motion tracking. An increase in frame rate, however, comes at the expense of a degraded resolution and contrast compared with conventional focused imaging. An approach to improve the quality of ultrafast imaging is coherent compounding [11], [12], where consecutive steered transmit beams at different angles are transmitted and the backscattered echoes are summed coherently. Coherent compounding can yield images of high quality, at the expense of the frame rate compared with single plane-wave/diverging-wave imaging. In practice, a trade-off must be made between image quality and frame rate.

In cardiac imaging, which involves fast-moving tissues, motion artifacts may ensue during a large number of transmit-receive events for coherent compounding, which in turn degrades the quality of compounded images. As a result, inter-acquisition motion must be considered during the compounding process to mitigate image degradation. Motion compensation (MoCo) methods have thus been introduced to provide high-quality echocardiography [13], while preserving the speckle patterns for carrying out efficient speckle tracking [14]. A typical frame rate of MoCo-based compounding approaches for high-quality echocardiography in adults is around 250 frames/second, and is still limited by the temporal-spatial resolution trade-off. Yet, MoCo techniques can be impaired by the limited quality of the reference frame in registration-based MoCo [15], or by aliasing due to the Nyquist limit in Doppler-based MoCo [13]. Exploiting the full potential of ultrafast cardiac imaging in terms of frame rate and image quality is a very active area of interest.

Recently, a wide variety of studies based on deep learning have sought to improve the image quality of high-frame-rate ultrasound imaging, such as PW imaging [16]–[23], DW imaging [24]–[26], synthetic transmit aperture imaging [27], [28], multiline acquisition [29], [30], and multiline transmission [31]. Among these works, the studies in [29]–[31] were devoted to cardiac imaging. In these studies, the image quality was typically only evaluated on static cardiac frames in terms of image contrast, as well as similarity to high-quality references from single-line acquisitions. However, the interest of high-frame-rate echocardiography lies in deciphering highly transient physiological events using cardiac sequences in high temporal resolution. Deep learning-based methods have yet to show that it is possible to achieve high imaging quality while maintaining the capability of tracking blood and tissue motion.

However, very few studies have investigated the properties of deep learning-based reconstruction for motion estimation. One notable exception is the study in [21], which performed such investigation in the context of vascular imaging using PW imaging. In this work, motion estimation assessment was performed from numerical simulations. The presented *in vivo* results are obtained from one proof-of-concept case acquired using a low frame rate (10 Hz) and concerned displacements of slow-moving tissue between the skin and the carotid (i.e. 1.25 mm/s), which are thus far from the much more complex and fast cardiac motion patterns.

DW transmissions are particularly interesting in echocardiography, since they can cover a wide and deep cardiac view (e.g. a four-chamber view) with a single transmission [5], [14]. However, very few studies [24], [32], [33] are devoted to deep learning-based processing of DW. We previously described a deep learning framework for high-quality ultrafast DW imaging reconstruction [25], [26]. In [26], a complex-values CNN was exploited for reconstructing high-quality ultrasound images from DAS beamformed I/Q signals acquired with a small number (typically three) of DW emissions.

In this paper, we propose a deep learning-based reconstruction approach for ultrafast cardiac imaging, and showed that it yielded high imaging quality while achieving a high accuracy in tracking tissue motion.

The main contributions of this work are the following:

1) We designed a specific numerical simulation pipeline providing clinical-like ultrasound cardiac sequences, for proper training of deep learning models and assessing reconstruction and motion estimation quality. In particular, this simulation provided task-specific reference data, by allowing imaging the heart as if it was frozen in time, yielding high-quality, motion artifact-free images from the compounding of 31 DWs.

2) This simulation framework allowed to demonstrate that the proposed method reconstructs high-quality cardiac images using only 3 DWs, yielding an image quality and a motion accuracy equivalent to those obtained from the reference data, i.e. from the compounding of 31 motion artifact-free DWs.

3) The CNN trained on the cardiac simulation and the resulting motion estimation were further evaluated on non-simulated, experimental *in vitro* data, using a spinning disk phantom. This experiment allowed to demonstrate that our approach yields high-quality image reconstruction and accurate motion estimation under a large range of velocities. In particular, we showed that the proposed approach outperforms a state-of-the-art MoCo-based approach at high velocities.

4) The approach was assessed from *in vivo* cardiac acquisitions and the results showed consistent improvement in

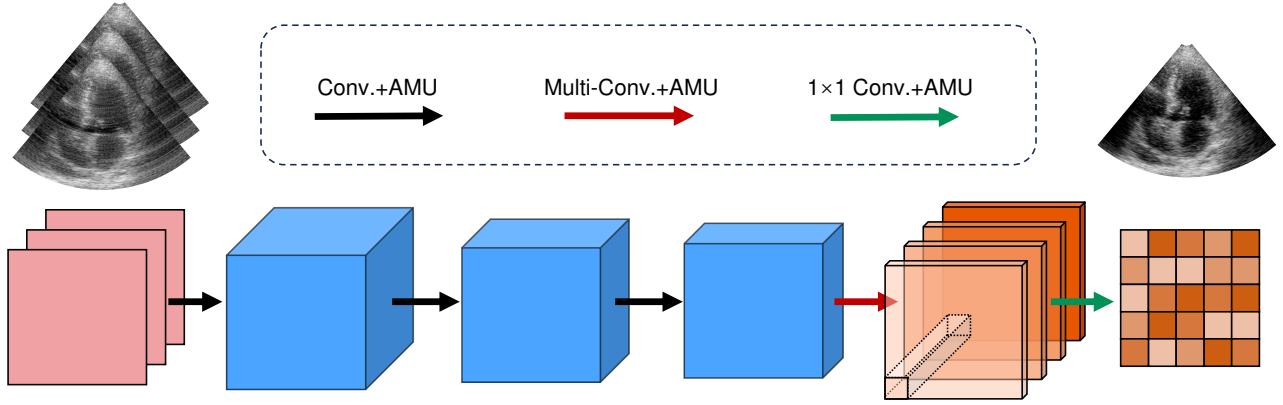


Fig. 1. Block diagram of the model architecture

TABLE I  
ARCHITECTURE OF THE PROPOSED NETWORK

Feature size	Kernel size	Kernel number	Activation
$m \times h \times w$	–	–	–
$64 \times h \times w$	$3 \times 3$	256	4-piece AMU
$32 \times h \times w$	$5 \times 5$	128	4-piece AMU
$16 \times h \times w$	$9 \times 9$	64	4-piece AMU
$8 \times h \times w$	$11 \times 11$	8	4-piece AMU
	$13 \times 13$	8	4-piece AMU
	$15 \times 15$	8	4-piece AMU
	$17 \times 17$	8	4-piece AMU
$1 \times h \times w$	$1 \times 1$	4	4-piece AMU

imaging quality and motion estimation, as compared with standard compounding, which demonstrated the feasibility and effectiveness of the proposed approach for ultrafast speckle-tracking echocardiography.

The remainder of this article is organized as follows. In Section II, we present our framework for ultrafast DW imaging, data simulation, and speckle tracking. The experimental configurations for realistic imaging and model training are presented in Section III. Section IV presents the experimental results, which are further discussed in Section V. Finally, our concluding remarks are given in Section VI.

## II. METHODS

### A. Deep Learning-based DW imaging

We consider a small number of  $m$  steered DW transmissions, each yielding raw I/Q image  $\chi_i$ ,  $i \in [1, m]$ . A DAS beamformer was first applied to produce beamformed I/Q data  $\mathbf{x}_i = \text{DAS}(\chi_i)$  of size  $h \times w$ , which represents the dimension of beamforming grid. The second step consists of reconstructing a high-quality I/Q image from the set of beamformed data  $\mathbf{x} = \{\mathbf{x}_i\}$ . Instead of the standard compounding that requires a large number of transmits (typically on the order of thirty in cardiac imaging [13]), we adopted a CNN-based model with learnable parameters  $\theta$  to seek an optimal DW compounding  $\hat{\mathbf{y}} = f_{\theta}(\mathbf{x})$  using supervised learning, with respect to high-quality targets  $\mathbf{y}$ . The references in this study were obtained

from the compounding of a large number of motion artifact-free DWs. The model was trained by searching for the optimal convolutional weights  $\hat{\theta}$  with loss minimization,

$$\hat{\theta} = \arg \min \sum_{(x,y)}^{\mathcal{D}} \|f_{\theta}(x) - y\|_2^2 \quad (1)$$

where  $\mathcal{D}$  is a dataset containing representative  $(x, y)$  samples. In order to consistently process the I/Q data in the complex-valued domain, we adopted the complex-valued CNN model proposed in [26], briefly described hereafter. The interested reader will find in [25] and [26] more details about the justification and properties of this architecture.

Fig. 1 shows the block diagram of the model architecture and Table I provides a more detailed specification. As we use 3 DWs as input to the network, the input dimension  $m$  is fixed to 3. The model is a fully convolutional network with five complex-valued convolution layers and amplitude maxout units (AMUs). We excluded spatial pooling for preserving the spatial information at the same resolution throughout the network. Each convolution layer is followed by an AMU activation, which activates elements corresponding to the element-wise maximum amplitude across several affine feature maps.

The second-to-last layer consists of the concatenation of multi-scale convolution kernels.  $1 \times 1$  convolution is used in the last layer for channel dimension reduction and producing reconstructed I/Q images. This layer is equivalent to cross-channel parametric pooling, which can be trained to select element-wise the main feature that contributes to the production of the output element. As demonstrated in [25], the multi-scale convolution used in conjunction with  $1 \times 1$  convolution and maxout activation allowed position-dependent features from different receptive fields to be captured, which helped to address the specific geometry of data sampling for DW imaging.

### B. Cardiac Data Simulation for Deep Learning

Effective learning of the reconstruction model generally requires the dataset  $\mathcal{D}$  to be representative of the joint distribution of input-output space, i.e. paired  $(x, y)$  with a wide variety of cardiac geometries in the context of echocardiography. It is however challenging to obtain high-quality target images

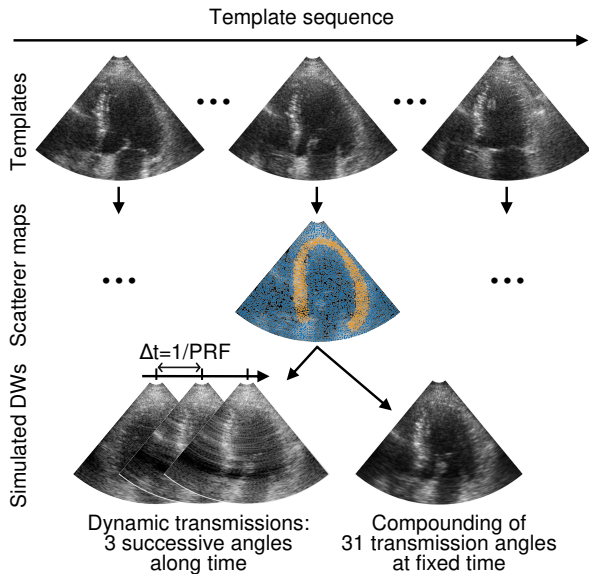


Fig. 2. Framework of cardiac imaging simulation. Based on real sequences, 3 DW transmissions were simulated with cardiac motion, and 31 DW transmissions were simulated at the frozen time.

$y$  with realistic acquisitions, since cardiac motion induces artifacts when a large number of transmit-receive events are used for coherent compounding. We therefore propose a simulation pipeline for generating clinical-like ultrasound cardiac sequences to tackle this issue. We briefly describe in the following the specificities of the simulation pipeline yielding the time-matched pairs  $(x, y)$ . The interested readers will find a more detailed description in [34], [35].

1) *Pipeline overview*: The overall framework for the numerical simulation is illustrated in Fig. 2. The pipeline starts with B-mode cardiac sequences from clinical recordings. These real images were used as the templates for cardiac geometries, which define the acoustic scatterer maps in the simulated field of view. Steered DW transmissions in the scatterer medium were simulated using a homemade open-source software called SIMUS [36], [37] for the generation of synthetic ultrasound data. To obtain motion artifact-free reference, each  $y$  was generated with steered DWs fired at a single time point (i.e. with a fixed scatterer map corresponding to the simulated frame), thus allowing imaging the heart as if it was frozen in time. On the opposite,  $x$  was produced in a dynamic manner for mimicking real acquisitions. The steered DW transmissions were simulated while taking into account the cardiac motion between each acquisition. The cardiac motion was simulated by the scatterer displacement using the strategy described at the end of this section.

2) *Cardiac geometry*: A carefully-defined medium of acoustic scatterers was exploited for the simulation of cardiac geometry. In order to obtain realistic speckle statistics, we followed the approach previously described in [34]. For each template sequence, the scatterers were randomly distributed within the imaging sector at a density of 10 per square wavelength and the reflection coefficients  $RC_m$  of the scatterers were computed according to the local intensities  $I_m$  of the B-mode templates

as [34]

$$RC_m = (I_m/255)^{(1/\gamma)} \cdot \mathcal{N}(0, 1) \quad (2)$$

where  $\mathcal{N}(\cdot)$  is the normal distribution, and  $\gamma$  is a constant for gamma compression. As shown in [34], this strategy yields fully developed speckle for blood regions (i.e. Rayleigh statistics) and partially developed speckle in myocardial regions (i.e. sub-Rayleigh statistics). The initial myocardium scatterers were distributed in the manually annotated myocardium region of the first frame. The scatterer positions for the subsequent frames were computed using the strategy described in Section II-B-3. The reflection coefficients of myocardium scatterers were set constant to maintain the speckle texture throughout the simulated sequences.

3) *Myocardium motion*: The simulation pipeline includes the motion fields in myocardium regions. The myocardium regions were obtained by manually delineating the endocardial and epicardial contours on the template frames. The regions of interest (ROIs) were then resampled to obtain time-varying surface meshes throughout the cardiac cycles. The myocardium motion fields were obtained with the propagation of the mesh over the full sequence. Such a scheme allows obtaining paired  $(x, y)$  at any time point throughout the temporal trajectory of the simulated motion. The motion fields in the simulation were used as the ground truth for the evaluation of speckle tracking among consecutive reconstructed frames.

### C. Speckle tracking

The tissue interframe motion was estimated on reconstructed frames using a speckle tracking approach described in [38]. This approach is based on a block-matching algorithm, where the normalized cross-correlation (NCC) is computed via a FFT for efficiency [39]. Let us note that other efficient approaches for NCC computation are possible [40]. The approach can be summarized by the following process. The consecutive real-envelope images were divided into overlapping subwindows. The normalized cross-correlation (NCC) corresponding to each subwindow was computed on the subwindow pairs of two successive frames, and the relative displacement within each subwindow is given by the location of the peak in each NCC. A parabolic peak fitting around the NCC peak was used to obtain the displacement with subpixel accuracy. To detect a wide range of motions, the process was iteratively performed in a coarse-to-fine manner by iteratively decreasing the size of the subwindows to refine the motion fields.

## III. EXPERIMENTS

### A. Data acquisition

1) *Simulated data*: We performed the cardiac sequence simulation using the pipeline described in Section II. The template sequences were extracted from the CAMUS open-access dataset [41] composed of exams from 500 patients acquired in clinical routine from the University Hospital of St-Etienne (France) within the regulation set by the local ethical committee of the hospital. A subset of 94 sequences with apical four-chamber views were selected for the simulation,

resulting in 4324 frame templates with the corresponding cardiac texture and myocardial displacement fields.

We simulated steered DW transmissions under the imaging settings corresponding to an ATL P4-2 probe: a 64-element phased array of 0.3-mm pitch and 50- $\mu\text{m}$  kerf. The center frequency was 3 MHz and the bandwidth was 2–4 MHz. For each template frame, we simulated 31 steered DW transmissions with tilt angles evenly spaced between  $\pm 20^\circ$ . The scatterer map was set fixed to obtain motion artifact-free references  $\mathbf{y}$ . We then simulated 3 steered DW transmissions  $\mathbf{x}$  while taking into account the cardiac motion between each transmission. The steering angles were  $-20^\circ$ ,  $0^\circ$ , and  $20^\circ$ . The motion was obtained by interpolating the reference motion field considering the pulse reception frequency (PRF), which was set to 3850 Hz in the simulation (i.e.  $\Delta t = 0.26$  ms). Such a PRF allows round-trip ultrasonic transmissions of 20-cm depth between each acquisition, which is the largest axial depth of the template sequences. The synthetic RF data were sampled at 12 MHz, then demodulated and beamformed using a delay-and-sum (DAS) [42] to produce the beamformed I/Q data.

2) *In vitro data*: The *in vitro* data were acquired using a research scanner (Vantage 256, Verasonics Inc.) and an ATL P4-2 probe with the same probe settings used for the simulation and a PRF of 4500 Hz. Imaging was carried out on a 10-cm-diameter tissue-mimicking disk phantom with four anechoic cysts. The disk was mounted on a step motor and rotated at controlled speeds. The angular velocities ranged from 0 to 12 radians per second with an increment of one, giving a maximum speed of 60 cm/s at the disk periphery that encompasses the myocardial maximum velocity, i.e.  $\sim 15$  cm/s for normal subjects [43] to  $\sim 30$  cm/s for endurance athletes [44]. We performed two separate DW acquisition series for each rotational speed: the 3-angle sequence and the MoCo-based one. The MoCo-based sequences used a triangle steering strategy with 32-angle series and 50% overlap [13]. This yields a frame rate of 280 frames/second for MoCo and 1500 frames/second for the 3-angle sequence. For each acquisition, the received raw RF data were sampled and processed with the same settings used for the simulation.

3) *In vivo data*: The *in vivo* study was performed from five normal subjects and was approved by the Polytechnique Montréal Comité d'éthique et de la recherche (CER-2122-54-D). Apical four-chamber views were examined. Two types of transmission, i.e. 3-angle sequences (used for CNN reconstruction) and 32-angle triangular sequences (used for MoCo-based reconstruction), were obtained with the same imaging and processing configurations used for the *in vitro* experiment. During acquisition, these two transmission types were interleaved, in order to minimize the time interval between the corresponding reconstructions and make them comparable. Using a PRF of 4500 Hz and considering a maximum myocardial velocity of 15 cm/s in normal subjects, this indeed corresponds to a maximum displacement smaller than 0.07 mm, i.e. less than  $\lambda/7$  (where  $\lambda$  denote the wavelength using a central frequency of 3 MHz and a sound speed of 1540 m/s).

## B. Training implementation

From the 94 simulated cases, 56 cases were randomly selected for training, and 19 cases were used as an independent validation set during training, to avoid overfitting. The remaining 19 cases were used for testing. In the training stage, the model weights were initialized with the Xavier initializer [45] and optimized by minimizing the training loss with mini-batch gradient descent. The batch size was set to 16 and the Adam [46] optimizer was employed for gradient descent with back-propagation. The initial learning rate was set to  $1 \times 10^{-4}$  and an early stopping strategy was used to adjust the learning rate. The learning rate was halved if the validation loss had not reduced for 10 epochs and the training ended if there had been 20 epochs without any validation loss reduction. The model was implemented using Pytorch library on an NVIDIA Tesla V100 GPU with 32 GB of memory, resulting in training time of about one day.

## C. Comparison Method

For numerical experiments, the results obtained from the motion artifact-free sequences (i.e. compounding of 31 DWs simulated at frozen time) were considered as the reference. The CNN-based results were then compared with this reference and with the results obtained from standard compounding of 3 simulated DWs (which took into account the cardiac motion between each transmission).

For *in vitro* and *in vivo* experiments, as a motion artifact-free reference is not available, the results obtained from a state-of-the-art MoCo-based compounding [13] were used as a reference. The CNN-based results were then compared with this reference and with the results obtained from standard compounding of 3 DWs.

## D. Evaluation Metrics

1) *Image quality*: We evaluated the image quality using peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) in the numerical experiment. These two metrics measure the image similarity between the reconstructed images  $\hat{\mathbf{y}}$  and the high-quality references  $\mathbf{y}$ .

$$\text{PSNR} = 20 \log_{10} \frac{\text{MAX}(\mathbf{y})}{\sqrt{\frac{1}{n} \|\hat{\mathbf{y}} - \mathbf{y}\|_2^2}} \quad (3)$$

where  $n$  is the number of pixels in the images.

$$\text{SSIM} = \frac{(2\mu_{\hat{\mathbf{y}}}\mu_{\mathbf{y}} + C_1)(2\sigma_{\hat{\mathbf{y}}\mathbf{y}} + C_2)}{(\mu_{\hat{\mathbf{y}}}^2 + \mu_{\mathbf{y}}^2 + C_1)(\sigma_{\hat{\mathbf{y}}}^2 + \sigma_{\mathbf{y}}^2 + C_2)} \quad (4)$$

where  $\mu_{\hat{\mathbf{y}}}$  and  $\mu_{\mathbf{y}}$  ( $\sigma_{\hat{\mathbf{y}}}^2$  and  $\sigma_{\mathbf{y}}^2$ ) are the means (variances) of  $\hat{\mathbf{y}}$  and  $\mathbf{y}$ , respectively,  $\sigma_{\hat{\mathbf{y}}\mathbf{y}}$  is the covariance between  $\hat{\mathbf{y}}$  and  $\mathbf{y}$ , and  $C_1$  and  $C_2$  are two constants that stabilize the division with a weak denominator.

For the *in vitro* experiments, we measured the contrast-to-noise ratio (CNR) and generalized contrast-to-noise ratio (gCNR) of the anechoic cysts in the disk phantom.

$$\text{CNR} = 20 \log_{10} \frac{|\mu_c - \mu_b|}{\sqrt{\sigma_c^2 + \sigma_b^2}} \quad (5)$$

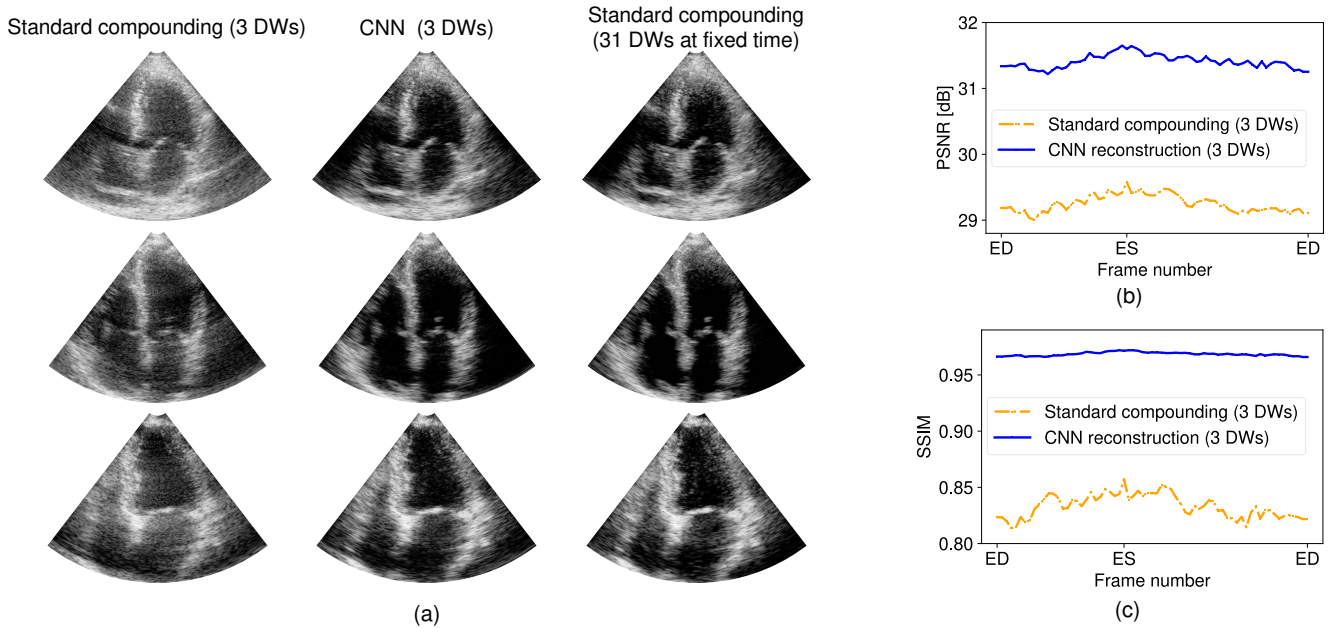


Fig. 3. (a) Typical reconstructions using standard compounding and CNN reconstruction from three DWs, along with the artifact-free reference reconstruction (from left to right) for 3 examples of cardiac simulations (from top to bottom). The images are displayed with a dynamic of 60 dB. (b) Average PSNR and (c) SSIM obtained using standard compounding and CNN reconstruction from three DWs for one cardiac cycle. For each time point, the average was calculated over the entire test set.

where  $\mu_c$  and  $\mu_b$  ( $\sigma_c^2$  and  $\sigma_b^2$ ) are the means (variances) of the intensities within the cyst and the background regions, respectively.

The gCNR is defined as the non-overlapping proportion between the intensity distribution within two regions.

$$\text{gCNR} = 1 - \sum_k \min\{p_c(k), p_b(k)\} \quad (6)$$

where  $k$  is the pixel intensity, and  $p_c(k)$  and  $p_b(k)$  are the probability mass functions of the pixel intensity  $k$  within the cyst and background regions.

2) *Motion estimation*: The accuracy of the estimated inter-frame motion was evaluated with the endpoint error (EPE), mean endpoint error (MEPE), and relative angular velocity error (RAVE). Given an estimated motion vector  $\hat{\mathbf{m}}$  and the true motion vector  $\mathbf{m}$ , the EPE is computed as the Euclidean distance between  $\hat{\mathbf{m}}$  and  $\mathbf{m}$ ,

$$\text{EPE} = \|\hat{\mathbf{m}} - \mathbf{m}\|_2 \quad (7)$$

For an ROI including  $n$  measure points, the MEPE is computed as the average of all EPE values over the set.

$$\text{MEPE} = \frac{1}{n} \sum_{i=1}^n \|\hat{\mathbf{m}}_i - \mathbf{m}_i\|_2 \quad (8)$$

The RAVE was specifically used for the *in vitro* experiment with the spinning disk.

$$\text{RAVE} = \frac{|\hat{\omega} - \omega|}{\omega} \quad (9)$$

where  $\omega$  is the true angular velocity controlled by the motor,  $\hat{\omega}$  is the estimated angular velocity, which was computed as the average of angular velocity at each inspected point within ROI covered by the spinning disk.

## IV. RESULTS

### A. Numerical Experiment

1) *Image quality*: Fig. 3(a) displays three examples of reconstructed cardiac images in the numerical experiments. One can observe that the CNN improved the contrast and enhanced the myocardium textures, as compared with the images obtained from the standard compounding of the same three DWs. In particular, the images produced by the CNN are visually close to the reference images, i.e. motion artifact-free images generated with 31 steered DWs fired at a fixed time.

We show in Fig. 3(b) and 3(c) the average PSNR and SSIM obtained using standard compounding and CNN reconstruction from three DWs. The results are given for one cardiac cycle, by synchronizing the sequences to their end-diastolic (ED) and end-systolic (ES) frames. For each time point, the average was calculated over the entire test set. From Fig. 3 (b) and 3 (c), the time variation of PSNR and SSIM associated with the two approaches showed a similar tendency: the values increased from ED to ES frame and then decreased from ES to ED frame. The CNN reconstruction showed less fluctuation along the cardiac cycle (average PSNR and SSIM), as compared to the standard compounding. This is particularly true for the SSIM results.

Table II reports the average PSNR and SSIM, as well as their standard deviations on all testing images. The standard compounding reached 29.28 dB in PSNR and 0.84 in SSIM. The CNN reconstruction reached 31.44 dB in PSNR and 0.97 in SSIM, which implies that CNN reconstructions are closer to the references, i.e. motion artifact-free reference images.

2) *Motion estimation*: Fig. 4(a) shows a typical EPE spatial distribution obtained for one test case at ED and ES, as

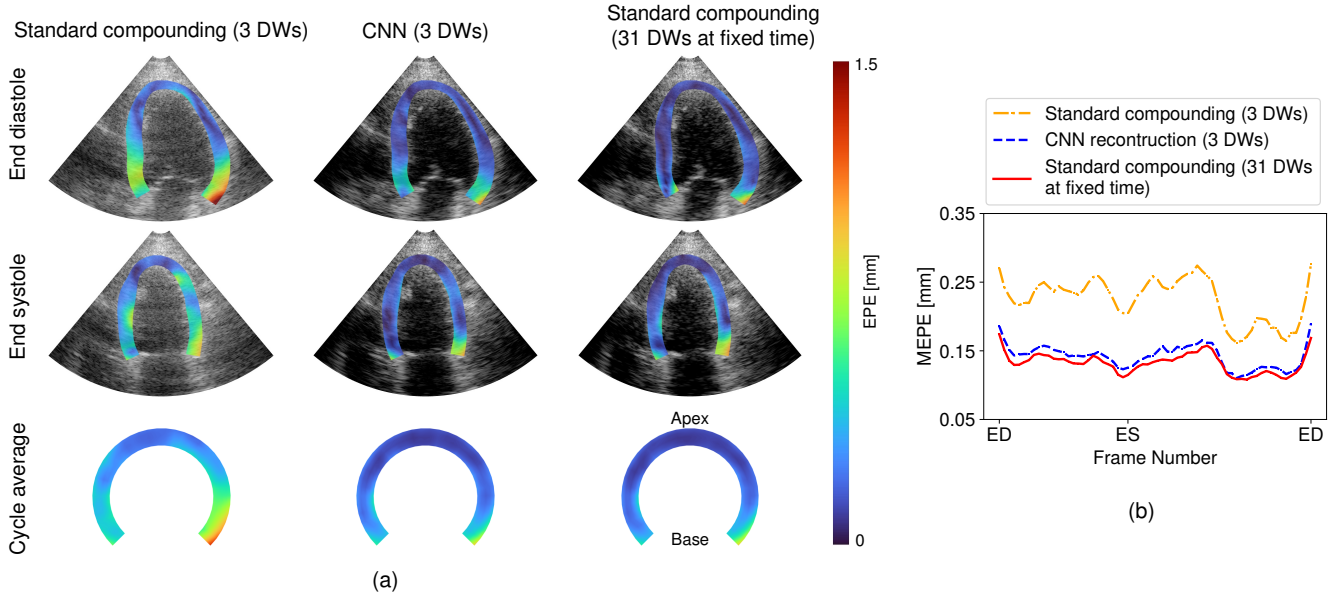


Fig. 4. (a) EPE spatial distribution of motion estimated from reconstructions using standard compounding and CNN reconstruction from three DWs, along with the result from the artifact-free reference (from left to right). Top and middle rows: typical EPE spatial distribution obtained for one test case at ED and ES. Bottom row: spatial distribution of the average EPE over the full cardiac cycle mapped on a common circular reference map for the same case. (b) Average EPE obtained using standard compounding, CNN reconstruction, and the motion artifact-free reference for one cardiac cycle. For each time point, the average was calculated over the entire test set.

TABLE II

PSNR, SSIM, AND MEPE REACHED BY DIFFERENT APPROACHES ON THE TESTING DATA.

Model	PSNR [dB]( $\uparrow$ )	SSIM( $\uparrow$ )	MEPE [mm]( $\downarrow$ )
Standard compounding (3 DWs)	$29.28 \pm 0.56$	$0.84 \pm 0.05$	$0.23 \pm 0.03$
CNN reconstruction (3 DWs)	$31.44 \pm 0.50$	$0.97 \pm 0.01$	$0.15 \pm 0.02$
Standard compounding (31 DWs at fixed time)	–	–	$0.13 \pm 0.01$

well as time-average distribution (averaged over one cardiac cycle). Due to the varying shape of the left ventricle along the cardiac cycle, the spatial distribution of the averaged EPE was mapped on a circular reference map. It appears that the CNN reconstruction leads an EPE distribution close to the one computed from the reference, and lower than the one associated with the standard compounding of 3 DWs.

Fig. 4(b) shows the time variation of the MEPE for each approach. For each time point, the MEPE average was calculated over the entire test set. One can observe that the three approaches showed a similar MEPE evolution tendency, while the CNN yielded a MEPE close to the one associated with the reference, and lower than the one associated with the standard compounding of 3 DWs throughout the whole cycle. Table II gives the MEPE for the three approaches, which were obtained by averaging the results on all testing data and time. Consistent with the observation in Fig. 4(b), the CNN reconstruction and reference reached rather close results, which were 0.15 mm and 0.13 mm, respectively. The standard compounding of 3 DWs yielded a higher MEPE, i.e. 0.23 mm.

## B. In Vitro Experiment

1) *Image quality*: Fig. 5(a) shows representative frames of reconstructed disk images in the *in vitro* experiment. From Fig. 5(a), it appears that the images obtained from the standard compounding of three DWs suffered from image artifacts, particularly for higher rotational speeds. Using the same three DW transmissions, the proposed CNN-based method restrained these artifacts, yielding a clear disk structure and enhanced contrast in the cyst regions. The MoCo-based imaging approach also reduced image artifacts. As the disk rotational speed increased, the MoCo-based approach failed to recover the disk structure in the high-speed regions. The reconstruction started to exhibit impaired structure on the disk rim at a rotational speed of 6 rad/s, corresponding to a rim linear speed of 30 cm/s.

We evaluated the quality of the reconstructed images in terms of CNR and gCNR, as shown in Fig. 5(b) and 5(c). The results were measured on the four anechoic cysts of the disk phantom. As the contrast varies with the position in the field of view, the investigated images for each approach at each rotational speed correspond to the frames where the phantom cysts are in the “cross” position, as shown in Fig. 5(a). With the disk rotational speed below 6 rad/s, the CNN-based imaging with 3 DWs obtained a CNR of  $\sim 7.5$  dB level and a gCNR of  $\sim 0.8$  level, which was close to MoCo-based imaging with 32 DWs and higher than those obtained by the standard compounding of 3 DWs ( $\sim 4.5$  dB in CNR and  $\sim 0.6$  in gCNR). With the rotational speed exceeding 6 rad/s, the CNR and gCNR reached by MoCo-based imaging decreased rapidly and dropped to its minimum of  $-3.2$  dB in CNR and 0.53 in gCNR. Similarly, the results associated with the standard compounding decreased with higher disk

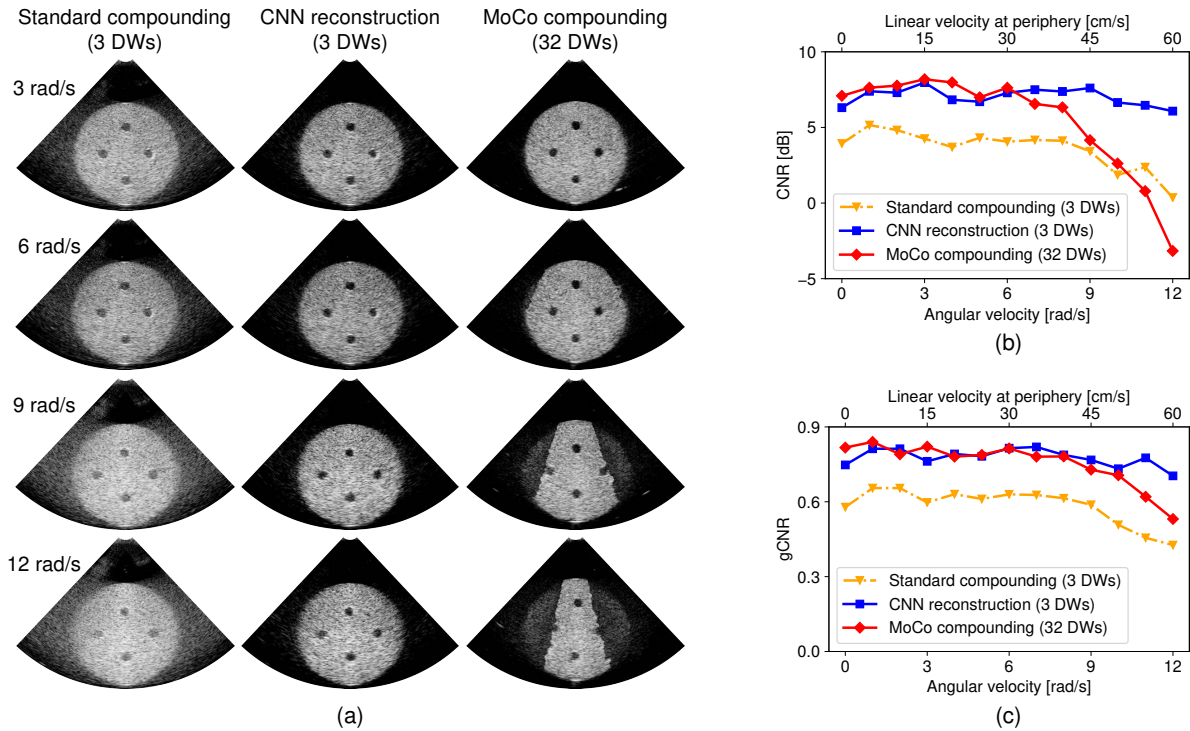


Fig. 5. (a) Reconstructed disk images at different rotational speeds, displayed in B-mode with a dynamic range of 60 dB. (b) CNR and (c) gCNR reached by different approaches at different disk rotational speeds. The CNR and gCNR were measured on the four anechoic cysts of the disk phantom at each rotational speed corresponding to the frames in (a).

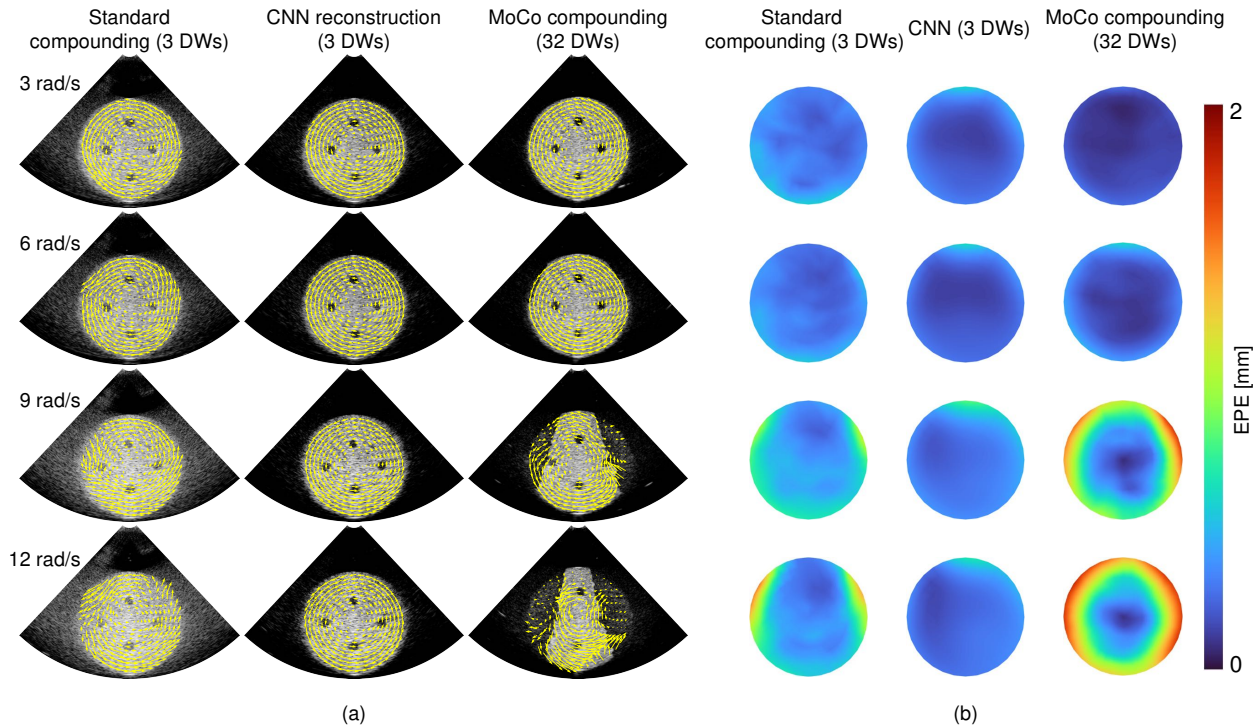


Fig. 6. (a) Motion fields and (b) EPE distribution of estimated motion on reconstructed disk sequences. Left to right: standard compounding of 3 DWs, CNN reconstruction of 3 DWs, and MoCo-based compounding of 32 DWs.

speeds, reaching its minimum of 0.36 dB in CNR and 0.43 in gCNR. The proposed CNN-based imaging yielded stable results (though slightly decreased) as the rotational speed

increased, and recovered a CNR of 6.6 dB and a gCNR of 0.72 at the highest rotational speed, resulting in a growing lead over the MoCo-based imaging and standard compounding. These



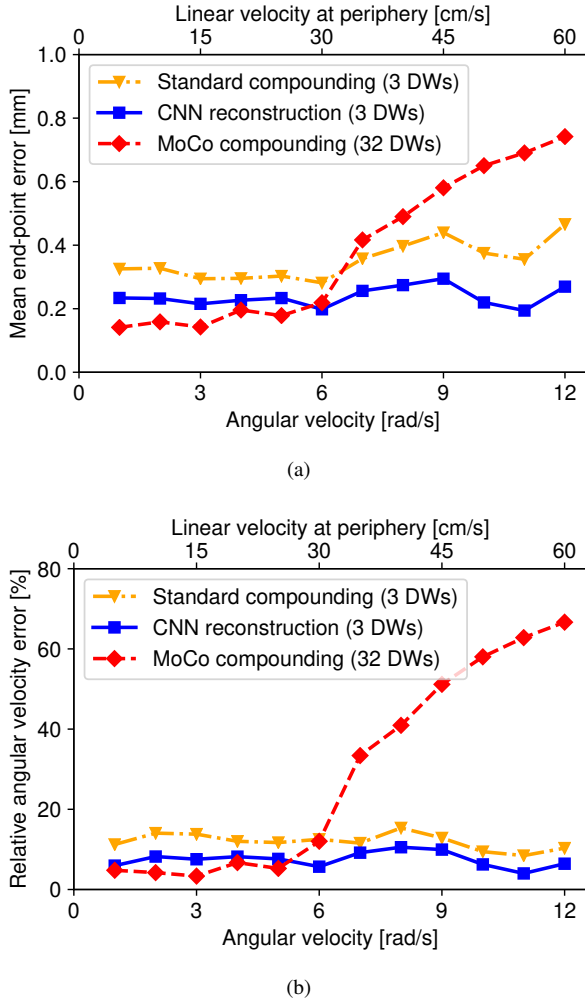


Fig. 7. (a) MEPE and (b) RAVE reached by different approaches at different disk rotational speeds.

results are consistent with the visual observation in Fig. 5(a)

2) *Motion estimation*: Fig. 6 shows the motion field estimated by the speckle tracking and the EPE distribution of the estimated motion. One can observe that the standard compounding-based tracking suffered from the motion artifacts in the reconstructed images, exhibiting disturbed motion estimation as the disk rotational speed increased. This is further demonstrated by the EPE distribution, particularly in the lateral region where the disk periphery was blurred due to the motion artifacts. Using the same three DW acquisitions, the proposed CNN-based approach provided a homogenous motion field in the ROI and yielded lower EPE for each disk rotational speed. In the same way, homogenous and accurate motion fields can be observed in the MoCo-based tracking with the disk spinning at no higher than 6 rad/s. As the disk rotational speed increased, the MoCo-based tracking failed to obtain homogenous motion fields due to the inconsistent speckle pattern in the impaired structure, yielding higher EPE in these regions.

Fig. 7 reports the MEPE and RAVE associated with each approach at each rotational speed. From Fig. 7(a), with the rotational speed below 6 rad/s, the MoCo-based tracking

returned a MEPE below 0.20 mm, which was lower than the other approaches. A significantly increased MEPE was returned by the MoCo-based tracking as the rotational speed increased, reaching 0.74 mm at the highest rotational speed. The proposed CNN-based approach led a consistent MEPE in the range of 0.20 to 0.27 mm, which was lower than the standard compounding-based tracking (0.28 to 0.47 mm), regardless of the disk rotational speed. In the same way, the proposed CNN-based approach returned a consistent RAVE ranging from 4% to 10% under a large range of rotational speed, which was lower than the standard compounding-based tracking (8% to 15%). The MoCo-based tracking led the lowest RAVE of  $\sim 5\%$  with the rotational speed below 6 rad/s, while reaching up to 68% as the rotational speed increased.

### C. In Vivo Experiment

Fig. 8 (a) displays one example of reconstructed *in vivo* cardiac images at end-systole (first row) and left ventricular motion fields during systole (second row) and diastole (third row). It appears that CNN-based reconstruction improved the contrast and enhanced anatomical structures as compared to the standard compounding of the same three DWs, and yielded an image quality close to the MoCo-based reconstruction. Furthermore, the motion fields associated with CNN-based and MoCo-based tracking showed very similar directions and amplitudes, while standard compounding-based tracking failed to provide homogenous motion vectors.

For each frame, the image quality and motion estimation associated to the 3-angle-based imaging (i.e. standard compounding and CNN-based reconstruction) were compared to the results provided by the 32-angle MoCo-based imaging (considered as the reference) using SSIM and mean endpoint difference (MEPD). As opposed to the EPE measure, the MEPD is only a comparison measure, since the true cardiac motion is unknown. The equation of the MEPD is thus the following:

$$\text{MEPD} = \frac{1}{n} \sum_{i=1}^n \|\hat{\mathbf{m}}_i^{\text{CNN}} - \hat{\mathbf{m}}_i^{\text{MoCo}}\|_2 \quad (10)$$

where  $n$  is the number of measure points,  $\hat{\mathbf{m}}_i^{\text{CNN}}$  and  $\hat{\mathbf{m}}_i^{\text{MoCo}}$  are the motion vectors estimated from the CNN reconstruction and MoCo reconstruction at measure point  $i$ .

Fig. 8 (b) and (c) show the time variation of the SSIM and MEPD. For each time point, the results were averaged over the five *in vivo* cases in one cardiac cycle. The CNN leads to a clearly higher SSIM than that associated with standard compounding, regardless of the time in the cardiac cycle. Regarding motion estimation, the CNN yields a lower MEPD (with a maximum of 0.12 mm) and less fluctuation than standard compounding. Consistently, the difference between the two approaches narrows considerably in late diastole, when the cardiac motion is small.

## V. DISCUSSION

### A. Learning Ultrafast Cardiac Imaging

In this study, we introduced a methodology for learning ultrafast cardiac imaging. The imaging process was formulated

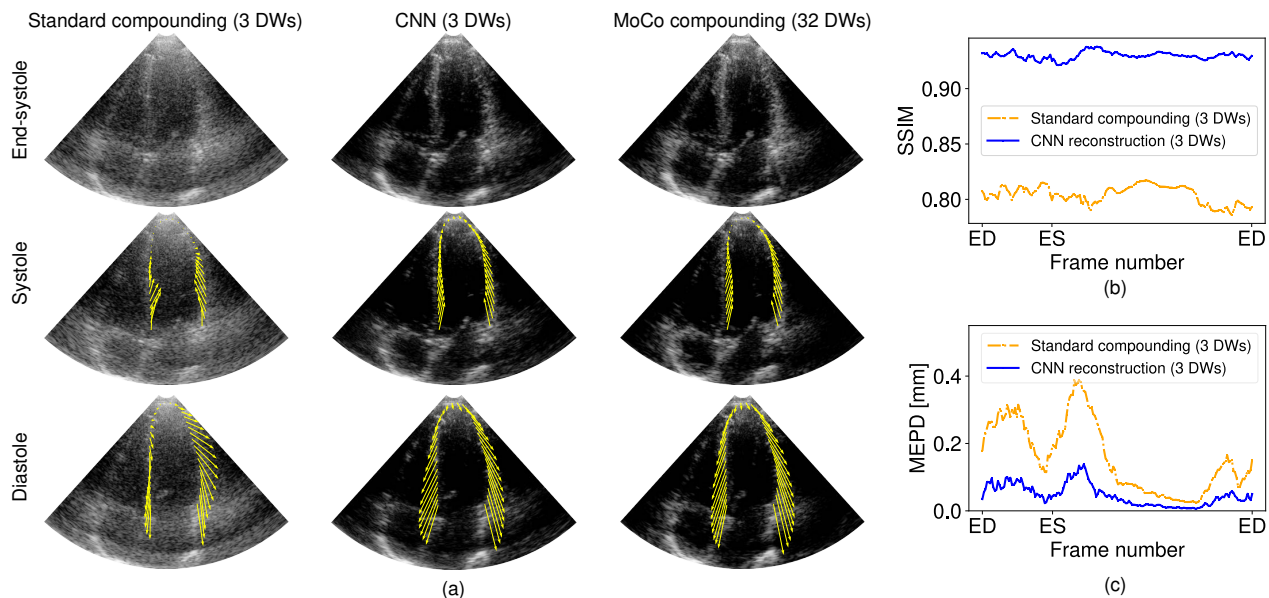


Fig. 8. (a) Reconstructed images at end-systole and left ventricular motion fields during systole and diastole for one example of *in vivo* case (from top to bottom), obtained from standard compounding and CNN reconstruction using three DWs, along with MoCo-based compounding from 32 DWs (from left to right). (b) Average SSIM and (c) Mean endpoint difference obtained using standard compounding and CNN reconstruction from three DWs for one cardiac cycle. For each time point, the average was calculated over the five *in vivo* cases.

as a non-linear mapping between the input space of low-quality images from ultrafast DW acquisitions to the output space of high-quality cardiac images, which was solved by supervised learning with respect to high-quality motion artifact-free references. Considering the dynamic nature of cardiac imaging, acquiring such reference is challenging, if not possible, under real conditions. We resolved this challenge by adopting a carefully-designed simulation pipeline. Taking advantage of numerical simulations, benchmark images can be pushed “beyond physics” by simulating optimal DW insonifications of a static heart medium to eliminate motion artifacts. Leveraging such a dataset, a complex-valued CNN underwent learning for seeking an optimal reconstruction operator, specifically trained to restore high-quality I/Q images from only three successive DW transmissions.

Visual assessment of the simulated cardiac images (Fig. 3 (a)) and the obtained metrics (Table II) demonstrated that the proposed approach significantly improved the image quality compared with standard compounding using the same three DW transmissions. *In vitro* and *in vivo* results (Fig. 5 and 8) showed consistent performance of the proposed method under real conditions, though the model was trained solely in the subspace of simulated cardiac data. Moreover, equivalent imaging quality was demonstrated in qualitative and quantitative assessments, as compared with motion artifact-free compounding in numerical experiments and MoCo-based compounding in physical experiments, which indicates a good generalization of the proposed approach under different imaging conditions.

Speckle tracking results in the numerical experiment associated with the proposed approach confirm the preservation of temporal consistency of tissue motion among successive reconstructed frames, with a MEPE very close to the reference throughout the cardiac cycle (only an increase of 0.02 mm in

average MEPE over the full cardiac cycle). *In vivo* cardiac results confirmed this trend, yielding a maximum mean endpoint difference of 0.12 mm between CNN and MoCo. This demonstrates that our CNN approach successfully learned an effective reconstruction operator for the underlying speckle pattern, toward a motion artifact-free quality.

These results demonstrate the feasibility of applying the deep learning-based method in ultrafast cardiac imaging, not only for improving exhibited image quality in separated frames but also preserving consistent speckle patterns in consecutive frames, which enables analysis of cardiac dynamics at ultrahigh frame rates.

### B. Imaging and Tracking on High-Velocity Tissues

Assessment of cardiac dynamics by speckle tracking requires high-quality imaging with consistent speckle patterns, which can be challenging when imaging and tracking high-velocity tissues. Ultrasound imaging with unfocused wavefronts can provide ultrahigh frame rates, but suffers from a low image quality if the transmission number for one frame is limited. As expected, the limitation was demonstrated with the *in vitro* experimental results associated with the standard compounding of 3 DWs, in terms of both imaging quality and speckle-tracking accuracy. Fig. 5 illustrates the poor visual quality of the reconstructed images from this approach. The quantitative assessment in terms of CNR and gCNR hold for this visual observation, yielding the lowest results (except for the CNR in the 11- and 12-rad/s cases which were higher than the MoCo-based method). Motion artifacts were observed in high-velocity cases from Fig. 5(a), leading to poorer speckle tracking in Fig. 6 and increased MEPE in Fig. 7.

As high-quality DW imaging generally relies on coherently compounding multiple steered transmissions, motion compen-

sation (MoCo) is required to deal with the inter-acquisition motion when imaging fast-moving tissues. The implemented MoCo-based method indeed demonstrated improved imaging quality in the *in vitro* experiment, as shown in Fig. 5(a) (0- and 3-rad/s cases), and yielded superior CNR and gCNR results when the rotational speed was not high. As the disk rotational speed increased above 6 rad/s, a performance deterioration in image quality and motion estimation was observed in Fig. 5(a) and 6. Such degradation can be demonstrated by the drop in CNR and gCNR, as well as the increase in MEPE and RAVE. The maximum detectable velocity of effective MoCo is related to its Doppler Nyquist limit, which can be derived by  $c PRF / (8 f_0)$  [13] (with  $f_0$  = carrier frequency). Considering a speed of sound  $c$  of 1540 m/s and the imaging settings in the experiments, the deployed MoCo reached a threshold of  $\sim 29$  cm/s, which is consistent with the observation in the 6-rad/s case (corresponding to 30 cm/s at the disk rim). For lower speed cases, MoCo is fully operational, while the CNN compensates slightly less for the motion intervening between each steered DW acquisition.

The interest of the present study was to improve image quality while preserving the temporal consistency of high-velocity tissue motion, thereby pushing the limit of detectable velocity for high-quality ultrasound imaging and accurate speckle tracking. An important observation from Fig. 5(a) and 6 is that the proposed method provided robust imaging quality and enabled accurate motion estimation under a large range of rotational velocities. Fig. 5(b), 5(c), and 7 further demonstrate steady performance in all evaluation metrics preserved by the proposed method, particularly at high velocities that exceed the limit of the MoCo-based imaging. Reliable speckle tracking was demonstrated in the 12-rad/s case, reaching the maximum velocity at the boundary with a value of 60 cm/s. Imaging modalities for higher velocities, such as stress echocardiography and blood flow tracking, may thus benefit from such detectable velocity.

The *in vivo* experiments aimed at exploring the feasibility and reliability of the proposed approach for future use in clinical settings, using a 32 DWs triangular MoCo sequence as a reference. The results (Fig. 8) were obtained from normal subjects and indicate that the CNN approach is on par with the MoCo approach in terms of motion estimation (maximum mean endpoint difference of 0.12 mm), while offering the possibility of much higher frame rates (1500 frames/second when using a PRF of 4500 Hz).

### C. Perspectives

The proposed approach provides a feasible solution for ultrafast speckle-tracking echocardiography, while relying on two separate steps, i.e. CNN-based image reconstruction from ultrafast DW transmissions and speckle tracking on reconstructed sequences. This process may be refined in an end-to-end manner by adopting multi-task learning for both imaging reconstruction and motion tracking, as deep learning methods have also been developed for motion estimation [35], [47]–[51]. The model accuracy may be further improved by considering joint optimization for multiple tasks and possible

motion compensation from the estimated motion. Moreover, as the proposed approach adopts data-driven deep learning, the training data is also a crucial factor for possible improvement. In the present study, the myocardial motions in the training data were limited to planar simulations. Integrating out-of-plane motions in the myocardium simulation may be interesting, as they generally occur in physical conditions and may bring additional speckle inconsistency in 2D. It would also be interesting to incorporate intracardiac flow (as recently proposed by [52]) in the training data for exploiting analysis diversity from the proposed ultrafast imaging, since DW-based echocardiography is well-adapted for simultaneous analysis of myocardial motion and intracardiac flow.

## VI. CONCLUSION

In this article, a methodology for ultrafast cardiac imaging based on deep learning was presented, particularly implemented in conjunction with speckle tracking for investigating its temporal property for motion estimation. The method consists of reconstructing high-quality cardiac images from ultrafast DW acquisitions by adopting a complex-valued CNN and supervised learning from a simulated echocardiographic dataset. We evaluated the performance of the proposed method in numerical, *in vitro*, and *in vivo* experiments, and demonstrated its effectiveness in both imaging quality enhancement and speckle pattern restoration for reliable speckle tracking, under a large range of tissue velocities, while allowing high frame rates. As such, cardiac diagnosis using dynamic analysis can benefit from the high imaging quality, frame rate, and perceptible speed of the proposed approach.

## REFERENCES

- [1] M. Cikes, L. Tong, G. R. Sutherland, and J. D’hooge, “Ultrafast cardiac ultrasound imaging: technical principles, applications, and clinical benefits,” *JACC: Cardiovascular Imaging*, vol. 7, no. 8, pp. 812–823, 2014.
- [2] J.-U. Voigt, G. Pedrizzetti, P. Lysyansky, T. H. Marwick, H. Houle, R. Baumann, S. Pedri, Y. Ito, Y. Abe, S. Metz *et al.*, “Definitions for a common standard for 2D speckle tracking echocardiography: consensus document of the eacvi/ase/industry task force to standardize deformation imaging,” *European Heart Journal-Cardiovascular Imaging*, vol. 16, no. 1, pp. 1–11, 2015.
- [3] P. A. Pellikka, A. Arruda-Olson, F. A. Chaudhry, M. H. Chen, J. E. Marshall, T. R. Porter, and S. G. Sawada, “Guidelines for performance, interpretation, and application of stress echocardiography in ischemic heart disease: from the american society of echocardiography,” *Journal of the American Society of Echocardiography*, vol. 33, no. 1, pp. 1–41, 2020.
- [4] K. Addetia, M. Takeuchi, F. Maffessanti, Y. Nagata, J. Hamilton, V. Mor-Avi, and R. M. Lang, “Simultaneous longitudinal strain in all 4 cardiac chambers: a novel method for comprehensive functional assessment of the heart,” *Circulation: Cardiovascular Imaging*, vol. 9, no. 3, p. e003895, 2016.
- [5] P. Joos, J. Porée, H. Liebgott, D. Vray, M. Baudet, J. Faurie, F. Tournoux, G. Cloutier, B. Nicolas, and D. Garcia, “High-frame-rate speckle-tracking echocardiography,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 65, no. 5, pp. 720–728, 2018.
- [6] J. Provost, W.-N. Lee, K. Fujikura, and E. E. Konofagou, “Electromechanical wave imaging of normal and ischemic hearts in vivo,” *IEEE transactions on medical imaging*, vol. 29, no. 3, pp. 625–635, 2009.
- [7] O. Villemain, M. Correia, E. Mousseaux, J. Baranger, S. Zarka, I. Podetti, G. Soulat, T. Damy, A. Hagege, M. Tanter *et al.*, “Myocardial stiffness evaluation using noninvasive shear wave imaging in healthy and hypertrophic cardiomyopathic adults,” *JACC: Cardiovascular Imaging*, vol. 12, no. 7 Part 1, pp. 1135–1145, 2019.

- [8] S. Salles, T. Espeland, A. Molares, S. A. Aase, T. A. Hammer, A. Støylen, S. Aakhus, L. Lovstakken, and H. Torp, "3D myocardial mechanical wave measurements: toward in vivo 3D myocardial elasticity mapping," *Cardiovascular Imaging*, vol. 14, no. 8, pp. 1495–1505, 2021.
- [9] L. Sandrin, S. Catheline, M. Tanter, X. Hennequin, and M. Fink, "Time-resolved pulsed elastography with ultrafast ultrasonic imaging," *Ultrasonic imaging*, vol. 21, no. 4, pp. 259–272, Oct. 1999.
- [10] H. Hasegawa and H. Kanai, "High-frame-rate echocardiography using diverging transmit beams and parallel receive beamforming," *Journal of medical ultrasonics*, vol. 38, no. 3, pp. 129–140, Jul. 2011.
- [11] G. Montaldo, M. Tanter, J. Bercoff, N. Bencech, and M. Fink, "Coherent plane-wave compounding for very high frame rate ultrasonography and transient elastography," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 56, no. 3, pp. 489–506, May. 2009.
- [12] C. Papadacci, M. Pernot, M. Couade, M. Fink, and M. Tanter, "High-contrast ultrafast imaging of the heart," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 61, no. 2, pp. 288–301, Feb. 2014.
- [13] J. Porée, D. Posada, A. Hodzic, F. Tournoux, G. Cloutier, and D. Garcia, "High-frame-rate echocardiography using coherent compounding with doppler-based motion-compensation," *IEEE Transactions on Medical Imaging*, vol. 35, no. 7, pp. 1647–1657, Jul. 2016.
- [14] J. Porée, M. Baudet, F. Tournoux, G. Cloutier, and D. Garcia, "A dual tissue-doppler optical-flow method for speckle tracking echocardiography at high frame rate," *IEEE Transactions on Medical Imaging*, vol. 37, no. 9, pp. 2022–2032, 2018.
- [15] A. Stanzola, M. Toulemonde, Y. Li, V. Papadopoulou, R. Corbett, N. Duncan, R. Eckersley, and M. Tang, "Motion artifacts and correction in multipulse high-frame rate contrast-enhanced ultrasound," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 66, no. 2, pp. 417–420, 2018.
- [16] M. Gasse, F. Millioz, E. Roux, D. Garcia, H. Liebgott, and D. Friboulet, "High-quality plane wave compounding using convolutional neural networks," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 64, no. 10, pp. 1637–1639, Oct. 2017.
- [17] A. A. Nair, T. D. Tran, A. Reiter, and M. A. L. Bell, "A deep learning based alternative to beamforming ultrasound images," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2018, pp. 3359–3363.
- [18] J. Lu and W. Liu, "Unsupervised super-resolution framework for medical ultrasound images using dilated convolutional neural networks," in *2018 IEEE 3rd International Conference on Image, Vision and Computing (ICIVC)*. IEEE, 2018, pp. 739–744.
- [19] S. Khan, J. Huh, and J. C. Ye, "Deep learning-based universal beamformer for ultrasound imaging," in *Medical Image Computing and Computer Assisted Intervention (MICCAI)*, 2019, pp. 619–627.
- [20] Z. Zhou, Y. Wang, Y. Guo, X. Jiang, and Y. Qi, "Ultrafast plane wave imaging with line-scan-quality using an ultrasound-transfer generative adversarial network," *IEEE journal of biomedical and health informatics*, vol. 24, no. 4, pp. 943–956, 2019.
- [21] D. Perdios, M. Vonlanthen, F. Martinez, M. Arditì, and J.-P. Thiran, "CNN-based ultrasound image reconstruction for ultrafast displacement tracking," *IEEE Transactions on Medical Imaging*, vol. 40, no. 3, pp. 1078–1089, 2020.
- [22] J. Zhang, Q. He, Y. Xiao, H. Zheng, C. Wang, and J. Luo, "Ultrasound image reconstruction from plane wave radio-frequency data by self-supervised deep neural network," *Medical Image Analysis*, vol. 70, p. 102018, 2021.
- [23] J.-Y. Lu, P.-Y. Lee, and C.-C. Huang, "Improving image quality for single-angle plane wave ultrasound imaging with convolutional neural network beamformer," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 69, no. 4, pp. 1326–1336, 2022.
- [24] M. U. Ghani, F. C. Meral, F. Vignon, and J.-I. Robert, "High frame-rate ultrasound imaging using deep learning beamforming," in *IEEE International Ultrasonics Symposium (IUS)*, 2019, pp. 295–298.
- [25] J. Lu, F. Millioz, D. Garcia, S. Salles, W. Liu, and D. Friboulet, "Reconstruction for diverging-wave imaging using deep convolutional neural networks," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 67, no. 12, pp. 2481–2492, Dec. 2020.
- [26] J. Lu, F. Millioz, D. Garcia, S. Salles, D. Ye, and D. Friboulet, "Complex convolutional neural networks for ultrafast ultrasound imaging reconstruction from in-phase/quadrature signal," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 69, no. 2, pp. 592–603, 2022.
- [27] Y. Chen, J. Liu, X. Luo, and J. Luo, "Apodnet: Learning for high frame rate synthetic transmit aperture ultrasound imaging," *IEEE Transactions on Medical Imaging*, vol. 40, no. 11, pp. 3190–3204, 2021.
- [28] J. Zhang, L. Huang, and J. Luo, "Deep null space learning improves dataset recovery for high frame rate synthetic transmit aperture imaging," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 70, no. 3, pp. 219–236, 2022.
- [29] O. Senouf, S. Vedula, G. Zurakhov, A. Bronstein, M. Zibulevsky, O. Michailovich, D. Adam, and D. Blondheim, "High frame-rate cardiac ultrasound imaging with deep learning," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2018, pp. 126–134.
- [30] S. Vedula, G. Senouf, A. Bronstein, O. Michailovich, and M. Zibulevsky, "Learning beamforming in ultrasound imaging," in *International Conference on Medical Imaging with Deep Learning*, 2019, pp. 493–511.
- [31] S. Vedula, O. Senouf, G. Zurakhov, A. Bronstein, M. Zibulevsky, O. Michailovich, D. Adam, and D. Gaitini, "High quality ultrasonic multi-line transmission through deep learning," in *International Workshop on Machine Learning for Medical Image Reconstruction*, 2018, pp. 147–155.
- [32] J. Lu, F. Millioz, D. Garcia, S. Salles, and D. Friboulet, "Fast diverging wave imaging using deep-learning-based compounding," in *2019 IEEE International Ultrasonics Symposium (IUS)*. IEEE, Oct. 2019, pp. 2341–2344.
- [33] J. Lu, F. Millioz, D. Garcia, S. Salles, D. Ye, and D. Friboulet, "Complex convolutional neural networks for fast diverging wave imaging," in *IEEE International Ultrasonics Symposium (IUS)*. IEEE, 2020, pp. 1–3.
- [34] M. Alessandrini, M. De Craene, O. Bernard, S. Giffard-Roisin, P. Allain, I. Waechter-Stehle, 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*, vol. 34, no. 7, pp. 1436–1451, 2015.
- [35] E. Evain, Y. Sun, K. Faraz, D. Garcia, E. Saloux, B. L. Gerber, M. De Craene, and O. Bernard, "Motion estimation by deep learning in 2D echocardiography: synthetic dataset and validation," *IEEE Transactions on Medical Imaging*, vol. 41, no. 8, pp. 1911–1924, 2022.
- [36] D. Garcia, "SIMUS: an open-source simulator for medical ultrasound imaging. Part I: theory & examples," *Computer Methods and Programs in Biomedicine*, vol. 218, p. 106726, 2022.
- [37] A. Cigier, F. Varray, and D. Garcia, "SIMUS: an open-source simulator for medical ultrasound imaging. Part II: comparison with four simulators," *Computer Methods and Programs in Biomedicine*, vol. 220, p. 106774, 2022.
- [38] V. Perrot and D. Garcia, "Back to basics in ultrasound velocimetry: tracking speckles by using a standard piv algorithm," in *IEEE International Ultrasonics Symposium (IUS)*, 2018, pp. 206–212.
- [39] D. Garcia, P. Lantelme, and E. Saloux, "Introduction to speckle tracking in cardiac ultrasound imaging," *Handbook of speckle filtering and tracking in cardiovascular ultrasound imaging and video. Institution of Engineering and Technology*, pp. 571–598, 2018.
- [40] J. Luo and E. E. Konofagou, "A fast normalized cross-correlation calculation method for motion estimation," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 57, no. 6, pp. 1347–1357, 2010.
- [41] S. Leclerc, E. Smistad, J. Pedrosa, A. Østvik, F. Cervenansky, F. Espinosa, T. Espeland, E. A. R. Berg, P.-M. Jodoin, T. Grenier *et al.*, "Deep learning for segmentation using an open large-scale dataset in 2D echocardiography," *IEEE Transactions on Medical Imaging*, vol. 38, no. 9, pp. 2198–2210, 2019.
- [42] V. Perrot, M. Polichetti, F. Varray, and D. Garcia, "So you think you can DAS? a viewpoint on delay-and-sum beamforming," *Ultrasonics*, vol. 111, p. 106309, 2021.
- [43] M. J. Garcia, L. Rodriguez, M. Ares, B. P. Griffin, J. D. Thomas, and A. L. Klein, "Differentiation of constrictive pericarditis from restrictive cardiomyopathy: assessment of left ventricular diastolic velocities in longitudinal axis by doppler tissue imaging," *Journal of the American College of Cardiology*, vol. 27, no. 1, pp. 108–114, 1996.
- [44] P. Caso, A. D'Andrea, M. Galderisi, B. Liccardo, S. Severino, L. De Simone, A. Izzo, L. D'Andrea, and N. Mininni, "Pulsed doppler tissue imaging in endurance athletes: relation between left ventricular preload and myocardial regional diastolic function," *The American Journal of Cardiology*, vol. 85, no. 9, pp. 1131–1136, 2000.
- [45] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, 2010, pp. 249–256.
- [46] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization,"

- in *Proceeding of International Conference on Learning Represent*, 2015, pp. 1–41.
- [47] A. K. Tehrani and H. Rivaz, “Displacement estimation in ultrasound elastography using pyramidal convolutional neural network,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 67, no. 12, pp. 2629–2639, 2020.
  - [48] R. Delaunay, Y. Hu, and T. Vercauteren, “An unsupervised learning approach to ultrasound strain elastography with spatio-temporal consistency,” *Physics in Medicine & Biology*, vol. 66, no. 17, p. 175031, 2021.
  - [49] X. Wei, Y. Wang, L. Ge, B. Peng, Q. He, R. Wang, L. Huang, Y. Xu, and J. Luo, “Unsupervised convolutional neural network for motion estimation in ultrasound elastography,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 69, no. 7, pp. 2236–2247, 2022.
  - [50] A. Østvik, I. M. Salte, E. Smistad, T. M. Nguyen, D. Melichova, H. Brunvand, K. Haugaa, T. Edvardsen, B. Grenne, and L. Lovstakken, “Myocardial function imaging in echocardiography using deep learning,” *IEEE Transactions on Medical Imaging*, vol. 40, no. 5, pp. 1340–1351, 2021.
  - [51] A. K. Tehrani, M. Ashikuzzaman, and H. Rivaz, “Lateral strain imaging using self-supervised and physically inspired constraints in unsupervised regularized elastography,” *IEEE Transactions on Medical Imaging*, 2022.
  - [52] Y. Sun, F. Vixege, K. Faraz, S. Mendez, F. Nicoud, D. Garcia, and O. Bernard, “A pipeline for the generation of synthetic cardiac color doppler,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 69, no. 3, pp. 932–941, 2022.