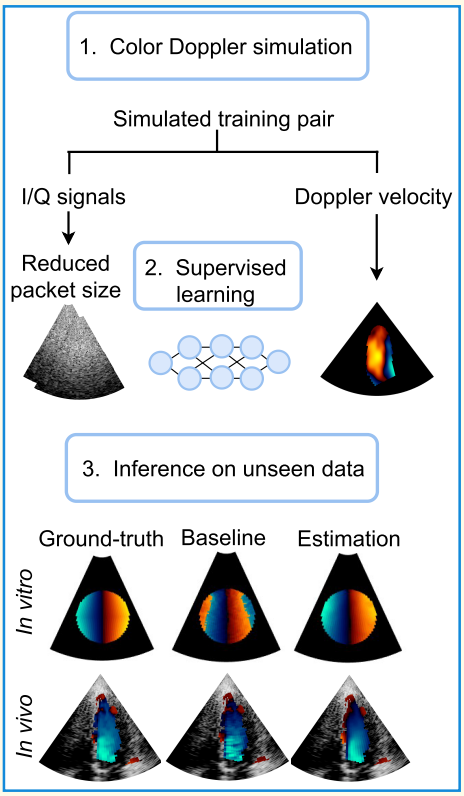
Boosting Cardiac Color Doppler Frame Rates With Deep Learning

Julia Puig, Denis Friboulet, Hang Jung Ling, François Varray, *Member, IEEE*, Michael Mougharbel, Jonathan Porée, Jean Provost, *Member, IEEE*, Damien Garcia, and Fabien Millioz

***Abstract*—Color Doppler echocardiography enables visualization of blood flow within the heart. However, the limited frame rate impedes the quantitative assessment of blood velocity throughout the cardiac cycle, thereby compromising a comprehensive analysis of ventricular filling. Concurrently, deep learning is demonstrating promising outcomes in postprocessing of echocardiographic data for various applications. This work explores the use of deep learning models for intracardiac Doppler velocity estimation from a reduced number of filtered I/Q signals. We used a supervised learning approach by simulating patient-based cardiac color Doppler acquisitions and proposed data augmentation strategies to enlarge the training dataset. We implemented architectures based on convolutional neural networks (CNNs). In particular, we focused on comparing the U-Net model and the recent ConvNeXt model, alongside assessing real-valued versus complex-valued representations. We found that both models outperformed the state-of-the-art autocorrelator method, effectively mitigating aliasing and noise. We did not observe significant differences between the use of real and complex data. Finally, we validated the models on in vitro and in vivo experiments. All models produced quantitatively comparable results to the baseline and were more robust to noise. ConvNeXt emerged as the sole model to achieve high-quality results on in vivo aliased samples. These results demonstrate the interest of supervised deep learning methods for Doppler velocity estimation from a reduced number of acquisitions.**



***Index Terms*—Aliasing mitigation, color Doppler, ConvNeXt, deep learning, echocardiography, ultrasound simulations.**

Manuscript received 31 May 2024; accepted 3 July 2024. Date of publication 8 July 2024; date of current version 27 November 2024. This work was supported in part by the Université de Lyon through the "Investissements d'Avenir" Program (ANR-11-IDEX-0007) operated by French National Research Agency (ANR) within the Framework of the LABEX PRIMES under Grant ANR-11-LABX-0063 and the LABEX CELYA under Grant ANR-11-LABX-0060. *(Corresponding author: Julia Puig.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethics and Research Committee of Polytechnique Montréal under Application CER-2122-54-D.

Julia Puig, Denis Friboulet, Hang Jung Ling, François Varray, Damien Garcia, and Fabien Millioz are with INSA-Lyon, Université Claude Bernard Lyon 1, CNRS, Inserm, CREATIS UMR 5220, U1294, F-69621 Lyon, France (e-mail: fabien.millioz@creatis.insa-lyon.fr; julia.puig@creatis.insa-lyon.fr).

Michael Mougharbel and Jonathan Porée are with the Department of Engineering Physics, Polytechnique Montréal, Montréal, QC H3T 1J4, Canada.

Jean Provost is with the Department of Engineering Physics, Polytechnique Montréal, Montreal, QC H3T 1J4, Canada, and also with Montreal Heart Institute, Montreal, QC H1T 1C8, Canada.

Digital Object Identifier 10.1109/TUFFC.2024.3424549

I. Introduction

COLOR Doppler echocardiography facilitates the concurrent visualization of cardiac tissue and intracardiac blood movement. In clinical settings, this imaging modality is extensively used as a visual aid to assess cardiac functions such as valvular regurgitation, ventricular filling, and ventricular diastolic function [1]. The frame rate achieved by focused color Doppler echocardiography is in the range of 10–30 frames/s assuming an apical long-axis acquisition from the apex to the basis of 12 cm and a number of lines between 20 and 60, depending on the sector width. Considering a standard subject with 80 bpm, this results in 8–23 frames per cardiac cycle. However, Krovetz and Goldbloom [2] found that around 25 harmonics are needed to accurately retrieve blood flow accelerations within the left ventricle, indicating that a minimum frame rate of 25 frames per cardiac cycle is necessary to quantitatively assess blood motion during all the phases of the cardiac cycle. Therefore, increasing the temporal

Highlights

- **We explored the use of deep learning for generating color Doppler images from clutter-filtered echocardiographic I/Q signals.**
- **We proposed a deep learning model that estimates phase shift from I/Q signals and mitigates aliasing from a reduced packet size. The model was trained solely on in silico data and validated on both in vitro and in vivo sequences.**
- **It was found that deep learning models can inherently reduce aliasing and are robust to noise induced by a reduced packet size.**

resolution of color Doppler could allow a better characterization of intracardiac blood movement and potentially open up new diagnosis possibilities. In particular, Faurie et al. [3] investigated the use of ultrafast imaging to achieve high frame rates of around 80 frames/s for color Doppler imaging. They found that high frame rates allowed to study blood vortex dynamics, and they derived quantitative tools that could be used to assess diastolic impairment. Here, we were interested in achieving high-frame-rate color Doppler in the conventional framework of focused imaging.

To generate a color Doppler image, backscattered echoes are acquired in the axial direction (fast time) at n consecutive instants (slow time). The radio frequency (RF) signals undergo I/Q demodulation, followed by clutter filtering and beamforming. The number n of slow-time acquisitions is referred to as the packet size. They enable the estimation of the phase shifts along the slow-time axis, from which Doppler velocities are subsequently derived. Phase shifts are usually estimated in two steps [4]. First, a clutter filter is applied to the I/Q signals to eliminate signals corresponding to tissues and slow-moving artifacts. Then, spatially weighted average autocorrelations are computed to estimate the local phase shifts [5]. In practice, the packet size typically ranges around $n = 8$. The reliability of the autocorrelation-based estimates significantly relies on the packet size; its reduction compromises robustness to noise and other artifacts.

In addition to noise, color Doppler can suffer from aliasing, which occurs when velocities exceed the Nyquist limit set by the pulse repetition frequency (PRF). While limited research has been dedicated to dealiasing in echocardiography, it is an essential step in extracting reliable quantitative information from color Doppler. In the context of ultrafast imaging, Posada et al. [6] proposed the use of staggered multi-PRF emissions to provide alias-free Doppler velocities. Deliasing has also been addressed as a postprocessing problem. Muth et al. [7] developed the dealiaser/denoiser (DeAN) algorithm for denoising and dealiasing color Doppler velocity maps. Their method involves unsupervised segmentation of aliased regions and subsequent dealiasing through comparison with neighboring regions. However, the method is prone to failure in cases of deteriorated signal quality.

In recent years, deep learning has led to significant advancements in medical imaging by replacing traditional model-based methods with learning-based approaches. In particular, convolutional neural networks (CNNs) have emerged as pivotal tools in ultrasound imaging, either by introducing entirely learning-based models or by enhancing the performance of existing model-based solutions with a deep learning step [8].

Few works have addressed Doppler velocity estimation with deep learning. For cardiac tissue Doppler estimation, van Sloun et al. [9] used an encoder–decoder on the I/Q signals to replace the classical autocorrelator method. Utilizing in vivo I/Q data from a porcine model, they derived Doppler estimates that achieved the quality of the autocorrelator method with reduced noise levels. In cardiac color Doppler, Apostolakis et al. [10] introduced a serial U-Net designed to estimate Doppler velocities from wall-filtered RF signals with reduced ($n = 4$) or undersampled packet sizes. They used a training dataset of flow phantom acquisitions whose reference velocities were given by an autocorrelator with a packet size of $n = 14$. He et al. [11] proposed a complex-valued CNN dedicated to carotid blood flow velocity estimation from RF signals, which simultaneously addressed clutter removal and velocity estimation, demonstrating promising results on simulated carotid data and in vivo experiments. Regarding clutter filtering, Solomon et al. [12] proposed a deep unfolding algorithm for super-resolution ultrasound applications. They formulated the problem as an iterative robust principal component analysis (RPCA) and solved it using a deep learning algorithm. In the same context, Brown et al. [13] performed clutter filtering on both in vitro and in vivo data using a spatiotemporal 3-D CNN.

Regarding dealiasing, Nahas et al. [14] trained a U-Net to identify and segment aliased regions in femoral bifurcation color Doppler images using Doppler frequency, power, and bandwidth information. Recently, they proposed an accelerated dealiasing technique for high-frame-rate vector Doppler imaging in femoral imaging [15]. Alternatively, Ling et al. [16] compared a deep unfolding method with a data-driven deep learning approach for dealing with aliasing in cardiac color Doppler images. The data-driven approach demonstrated superior results, with both methods outperforming the DeAN method mentioned above. All these dealiasing works can be considered as postprocessing methods applied to Doppler velocity maps.

In this study, we propose a deep learning approach as an alternative to the current autocorrelator technique in the context of focused cardiac color Doppler acquisitions. Our goal was to obtain accurate Doppler velocities from clutter-filtered I/Q signals with a reduced packet size of $n = 2$. We contended that CNNs can identify spatial and channel-wise connections

that the autocorrelator overlooks, potentially compensating for the packet size reduction. Our main contributions are given as follows.

1) We generated a training dataset for supervised learning by simulating color Doppler echocardiography with a dedicated pipeline. It included the generation of high-quality Doppler velocity ground truths and the simulation of realistic acquisitions of ultrasound signals.
2) We adapted state-of-the-art deep learning architectures for real-time Doppler velocity estimation from I/Q signals with a reduced packet size. We explored the performance of U-Net and ConvNeXt architectures, and the impact of using real-valued versus complex-valued representations.
3) We introduced a tailored *zooming* data augmentation procedure to increase the number of samples while ensuring signal coherence. In addition, we generated augmented *aliased* samples to make the models inherently robust to aliasing.
4) We validated the proposed deep learning models using both in vitro and in vivo sequences. These experiments showed that the deep learning models were able to generalize to domain shift data. They achieved comparable results to the autocorrelator while exhibiting robustness to noise and inherent reduction of aliasing.

II. Methods

A. Doppler Velocity Estimation in Color Doppler

In ultrasound imaging, the received RF signals undergo I/Q demodulation and beamforming. In the context of Doppler imaging, these resulting complex-valued signals can be organized in arrays of shape $h \times w \times n$, where h and w represent the fast-time and lateral dimensions, respectively, and n denotes the slow-time dimension referred to as the packet size. These processed signals are filtered to remove clutter and retrieve blood-related signals. Blood Doppler velocities are then estimated by calculating the phase shifts of the filtered signals in the slow-time direction, i.e., along the third dimension of the array. These phase shifts can be derived by evaluating the argument of the lag-one autocorrelator (R_1) applied to the I/Q signals. We adopted the 2-D autocorrelator introduced by Loupas et al. [5] as the baseline method. We smoothed R_1 spatially with a Hamming kernel of size 10×4 pixels to obtain a smoothed autocorrelation R_1^s. Finally, the Doppler velocities were computed by converting the arguments to velocities as follows:

$$v_D = -\frac{v_N}{\pi} \arg\left(R_1^s\right)$$

where v_N stands for the Nyquist velocity. The reliability of this technique depends on the packet size number, and decreasing it can degrade the estimation quality, especially when dealing with noisy acquisitions.

B. Color Doppler Simulations

Training supervised deep learning models requires large annotated datasets. In the context of color Doppler, high-quality in vivo ground truth is not available. Using directly the patient Doppler data, the models would learn to reproduce the behavior of the standard estimator along with its inaccuracies. Therefore, we decided to rely on a simulation approach that allows to generate a supervised training dataset with reliable ground truths. In this study, we adapted the patient-based color Doppler simulator developed by Sun et al. [17]. Unless otherwise specified, the parameters were consistent with those described in the original paper. A schematic representation of the inputs and outputs of the simulation pipeline is illustrated in Fig. 1. Given the patient's B-mode and Doppler information along a cardiac sequence, the simulation pipeline consisted of three main steps.

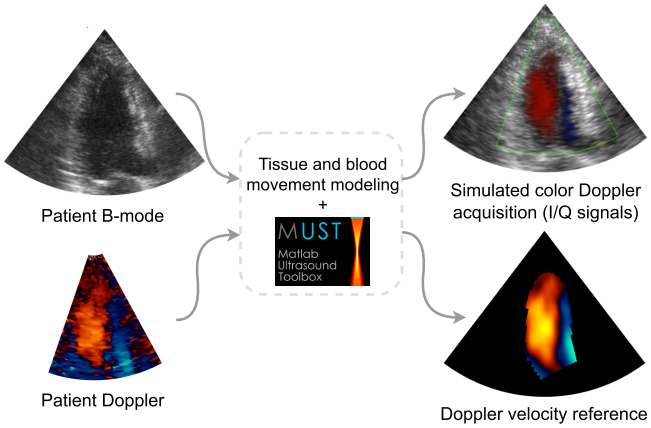


Fig. 1. Inputs and outputs of the color Doppler simulation pipeline.

1) Modeling Cardiac Tissue Motion: For each patient's B-mode image, the inner and outer walls of the left ventricular myocardium were segmented using a deep learning method [18]. A myocardial mesh was generated from the segmentation masks. Tissue scatterers were randomly distributed in the end-systole frame. Their positions were then propagated frame by frame based on the motion of the myocardial mesh. For additional information, refer to [17].

2) Modeling Blood Flow Motion: Patient Doppler velocity frames were treated as independent due to the low frame rate of color Doppler. For each color Doppler frame, a 2-D intraventricular vector flow map was computed from the patient's Doppler velocities using a physics-based method [19]. The method was used as described in the original paper. Then, for each frame, blood scatterers were distributed in the intraventricular region. A series of (number of firings $\times$ n) time steps was defined, during which the blood scatterers were displaced according to the 2-D velocity maps. To enhance realism, fluctuation values were computed from velocity maps and added to velocity values. This process was repeated for all frames of the sequence. From the resulting vector outputs, we used the radial components to build the ground-truth Doppler velocity maps $D \in \mathbb{R}^{h \times w}$ in our training dataset.

3) Simulation of an Ultrasound Acquisition: From the tissue and blood scatterer position maps defined at each time step, scanline-based ultrasound acquisitions were simulated using SIMUS [20], [21] of the MATLAB UltraSound Toolbox (MUST) [22]. The RF signals were postprocessed conventionally with I/Q demodulation and delay-and-sum beamforming on an $h \times w$ grid. The resulting clutter-free beamformed I/Q signals $S \in \mathbb{C}^{h \times w \times n}$ were the inputs in our training dataset.

C. Doppler Velocity Estimation With Deep Learning

We explored the use of CNNs, arguing that the spatial nature of convolutions could compensate for the loss of temporal information and produce better estimates. We focused on two architectures: U-Net and ConvNeXt.

1) U-Net: The U-Net architecture, introduced in 2015 for medical image segmentation, comprises an encoder–decoder structure with skip connections [23]. Given that our input S was complex-valued and the output D was real-valued, we examined two architectures: one entirely real-valued and another that was complex-valued in all layers except the last one.

1) *Real U-Net:* The input of the network was in the form $S \in \mathbb{R}^{h\times w\times 2n}$, where the real and imaginary parts of the I/Q signals were concatenated to obtain real representations. We began with the standard U-Net architecture as described in [23] and made two modifications. First, to ensure a lightweight network, we retained two dimension reduction steps, thereby reducing the network to only six layers. Second, the kernels of the two convolutions at each U-Net stage were of sizes 5×5 and 3×3. The number of feature maps was $(32, 64, 128)$ for the encoder part and symmetrically for the decoder part. A scheme of the architecture is shown in Fig. 2. With this configuration, the network had 1.5 million parameters, and the lowest dimension achieved was 45×10.
2) *Complex U-Net:* The input of the network was in the form $S \in \mathbb{C}^{h\times w\times n}$. The network features mirrored those of the Real U-Net, with the following distinctions. We employed complex-valued convolutions using the PyTorch framework. In particular, the network weights were complex-valued in all layers except the last one, where the real and imaginary parts of the feature map were concatenated before applying the real-valued weights. For the complex-valued activation function, we used the Complex ReLU, as advised by Trabelsi et al. [24]. Finally, we used the 2-D complex batch normalization available in [25], implemented following the recommendations of [24] in order to obtain equal variance in both the real and imaginary parts. The network contained the same number of parameters as the Real U-Net (1.5 M), but the convolutional operations were four times more computationally intensive.

2) ConvNeXt: We then considered ConvNeXt [26], a recent CNN architecture. Since 2020, hierarchical vision Transformer architectures have demonstrated superior performance over convolutional models in various computer vision tasks. However, Transformer models suffer from being data-intensive. As a solution, ConvNeXt was introduced in 2022. It is a convolutional-based model that integrates strategies from Transformer models, resulting in an architecture that outperforms state-of-the-art Transformers while remaining data efficient. One of its key features is the separate handling of spatial and channel information, akin to the self-attention mechanism in Transformers. This property is achieved through the use of depthwise separable convolutions. We adapted the ConvNeXt architecture to our problem by utilizing a U-Net architecture with ConvNeXt as its backbone (i.e., ConvNeXt layers served as feature extractors in the encoder part). A schematic representation of the considered architecture is depicted in Fig. 3. With this configuration, the network comprised 4.7 million parameters, and the lowest dimension attained was 11×5.

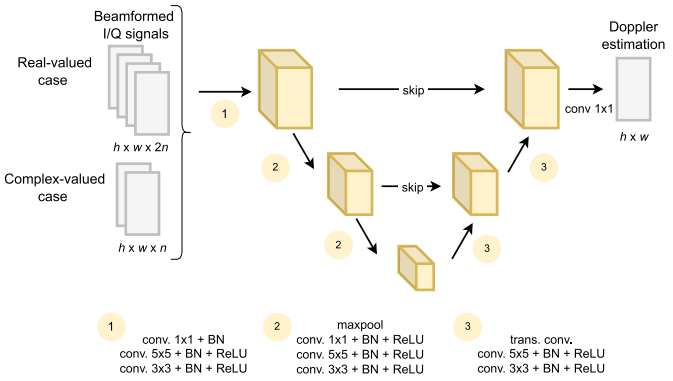


Fig. 2. Architecture of the U-Net-based model.

D. Data Augmentation

In computer vision, a common technique for data augmentation involves applying simple transformations to available images to increase the size and variability of the training set, resulting in a more robust model. However, for ultrasound data, the options for realistic data augmentation are limited due to the specific geometry of the acquisition process. In addition, some augmentation techniques require interpolation, which may significantly alter the signal information. Therefore, we only used vertical flipping and a custom zooming procedure.

1) Zoomed Samples: Standard simulations had $h \times w$ shape, where h and w refer to the original number of fast-time points and number of firings, respectively. Zooming was achieved by simulating color Doppler acquisitions with a finer grid than the one used for standard simulations and then cropping the result to an $h \times w$ shape at a random location. This was done by repeating simulations and increasing the number of fast-time points with a 1.5 ratio and the number of focused firings with a 1.5 ratio. The procedure is illustrated in Fig. 4.

2) Aliased Samples: Aliasing patterns appear in Doppler velocity maps when the maximum speed exceeds the Nyquist velocity. In such cases, phase unwrapping methods can be applied to the Doppler velocity maps as a postprocessing step, as described in Section I. In contrast to these postprocessing approaches, we aimed to make the Doppler estimation inherently robust to aliasing. This was achieved by providing the deep learning models with new training pairs (S, D), where S represents the I/Q aliased signal and D denotes the corresponding alias-free Doppler ground truth. The aliased samples were obtained during simulations by randomly reducing the original PRF values of each sequence by a factor in the $[0.4, 0.6]$ range. Each simulated aliased sample was considered as an *augmented* sample and added to the dataset.

III. Experiments

A. Data Acquisition

1) In Silico Experiment: Our simulations, described in Section II-B, aimed to neglect the influence of clutter through the exclusive modeling of tissue-free scenarios. As input,

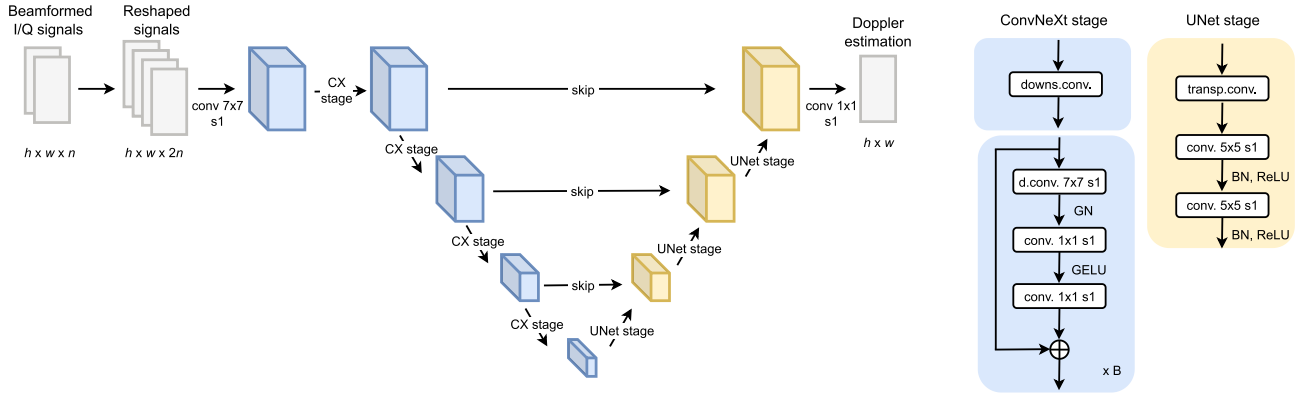


Fig. 3. Architecture of the ConvNeXt-based model. Initially, I/Q signals were reshaped by concatenating their real and imaginary parts. A first set of 7×7 convolutions was applied. Following this, four ConvNeXt stages were implemented, each containing $(3, 3, 9, 3)$ blocks, with a downsampling step in all stages except the first. The number of feature maps was $(32, 64, 128, 256)$ for the encoder part and symmetrically for the decoder part. "d.conv.," "trans.conv.," and "downs.conv." stand for depthwise, transposed, and pooling convolution, respectively. "s" indicates the convolution stride.

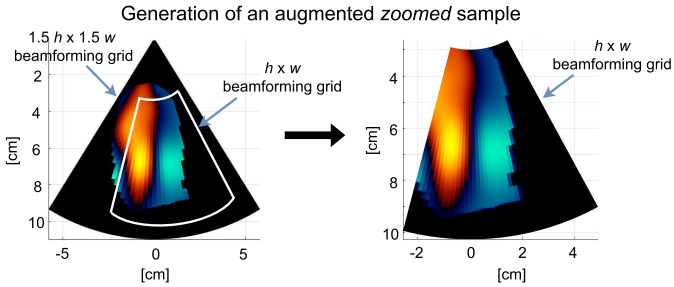


Fig. 4. Illustration of the zooming data augmentation procedure.

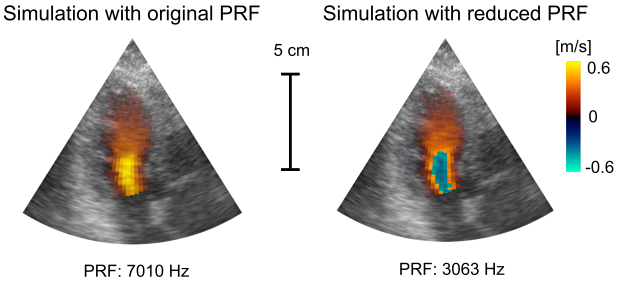


Fig. 5. Example of a simulated echocardiography acquisition with its original and decreased PRF. The Doppler velocities were estimated using an autocorrelator in both cases.

we used patient color Doppler echocardiographic data acquired in a previous study [27]. Examinations adhered to standard echocardiographic procedures, resulting in a diverse patient population encompassing individuals with and without cardiac disease. Due to randomization and anonymization, demographic and clinical details were unknown during the present work. The sequences, captured from the apical three-chamber view, incorporated both B-mode and Doppler velocity information. We obtained cardiac sequences of 37 patients, each containing at least one complete cardiac cycle (total: 2576 samples). From these patient echocardiographic data, we simulated 2576 training pairs. The probe and acquisition parameters are summarized in Table I. During the simulations, we leveraged patient-specific parameters such as PRF, sector angle, and depth. The simulated RF signals were demodulated into I/Q signals and beamformed into a 180×40 grid using delay-and-sum. Each training pair (S, D) consisted of I/Q signals $S \in \mathbb{C}^{180\times40\times n}$ and Doppler velocity references $D \in \mathbb{R}^{180\times40}$.

The simulated dataset is referred to as Original-Set and is described in Table II. A sample form Original-Set is shown in Fig. 5. Only 24 samples contained aliased regions, accounting for less than 0.001% of the total dataset pixels. This was circumvented with the simulation procedure described in Section II-D2 to increase the number of aliased samples by artificially decreasing the PRF. The set of aliased samples obtained is referred to as Aliased-Set in the sequel and is described in Table II. An example of a sample simulated with both its original and decreased PRF is depicted in Fig. 5. Finally, a third dataset was simulated to provide for augmented *zoomed* samples, as described in Section II-D1. This dataset is referred to as Zoomed-Set and is described in Table II.

From these three simulated datasets, we defined two training datasets: Train-Set-1, including Original-Set and nonaliased augmented samples of Zoomed-Set; and Train-Set-2, comprising all simulated datasets. They are described in Table III.

Finally, samples with reduced packet size were created by selecting two consecutive frames in the slow time.

TABLE I
PARAMETERS OF THE SIMULATED ULTRASOUND ACQUISITION

Parameter	Value
simulated probe	Verasonics P4-2v
central frequency (f_c)	2.7×10^6 Hz
pitch	300×10^{-6} m
#elements	64
bandwidth	74%
#cycles B-mode	1
#cycles Doppler	6
PRF	4500 - 7700 Hz
sector width	60° - 90°
depth	10 - 17 cm
sampling frequency	$4f_c$
acquisition type	focused waves
#firings B-mode	120
#firings Doppler	40
packet size (n)	8

TABLE II
DESCRIPTION OF THE THREE SIMULATED DATASETS

	Total	Non-aliased	Aliased
Original-Set	2,576	2,552	24
Aliased-Set	1,142	0	1,142
Zoomed-Set	2,576	1,883	693

TABLE III
DESCRIPTION OF THE TWO TRAINING DATASETS

	Source data	Total	Non-aliased	Aliased
Train-Set-1	Original-Set Zoomed-Set	4,459	4,435	24
Train-Set-2	Original-Set Aliased-Set Zoomed-Set	6,294	4,435	1,859

2) In Vitro Experiment: We designed a rotating disk experiment to measure Doppler velocities from a color Doppler acquisition. We constructed a 6-cm-diameter agar disk and connected it to a rotating motor. A Verasonics P4-2v cardiac probe was positioned at 8 cm from the center of the disk. To minimize clutter and artifacts, we submerged the entire setup in a water tank.

The acquisition parameters were set as described in Table I, with the sector width, depth, and PRF configured to 50°, 12 cm, and 6000 Hz, respectively. The packet size was $n = 32$. We ran a total of 12 experiments, varying the maximum outer speed between 0.1 and 1.13 m/s, mirroring intracardiac blood velocities. The Nyquist velocity was $v_N = 0.85$ m/s for all experiments, resulting in four experiments displaying aliased regions.

Samples with reduced packet size were created by selecting two consecutive frames in the slow time. The ground-truth Doppler velocity maps used to compute the models' estimation errors during inference were obtained from the known rotation speed of the phantom.

3) In Vivo Experiment: Cardiac echo-Doppler-focused acquisitions were conducted on six healthy volunteers. The study received approval from the ethics and research committee of Polytechnique Montréal (CER-2122-54-D). In accordance with the simulation and phantom experiments, the P4-2v Verasonics probe was used with identical parameters as those used for the phantom. Data were acquired in the apical three-chamber view, including both B-mode and Doppler information. The packet size was set to $n = 16$. The ultrasound signals were postprocessed and beamformed on a 180×40 grid. We obtained 8–16 color Doppler image samples from each volunteer.

To be consistent with the clutter-free setting of both in silico and in vitro experiments, we applied an singular value decomposition (SVD)-based clutter filter. We discarded the first four largest singular values to isolate the blood I/Q signals. We segmented the left ventricular endocardium to mask off I/Q signals not pertaining to intraventricular blood. Samples with low Doppler power in over 70% of the ventricular region were discarded, resulting in 50 selected samples. Only one of the selected samples contained aliased regions. To assess the method's robustness, we deliberately introduced aliasing by considering one out of two (or three) frames in the slow time, thereby artificially decreasing the PRF by a factor of 2 (or 3). We applied this procedure to the 50 selected samples, providing 20 new samples with several aliased regions.

Samples with reduced packet size were generated by selecting two consecutive frames in the slow time. In the in vivo experiments, ground-truth Doppler velocity maps were not available. Hence, Doppler velocity maps obtained with the autocorrelator using a packet size of $n = 8$ were used as references.

B. Deep Learning Implementation Details

1) Training Settings: Because the ultrasound data were simulated with patient-specific PRFs, the Nyquist velocity varied from sample to sample. Therefore, we opted to use phase shifts as the ground truth instead of Doppler velocities to ensure standardized inputs across samples. Each sample was normalized with respect to its maximum complex modulus across the slow time. For greater robustness, we trained the network using ninefold cross validation, where samples from a given patient were assigned to the same fold. Once the test fold was chosen, the split of samples in training/validation was set to 9/1 while ensuring that samples from a given patient remained in the same split. The network weights were initialized with the Xavier distribution. The loss function was the masked mean squared error, where the mask corresponded to the left ventricular cavity. The optimizer was AdamW, and the batch size was set to 16. The initial value of the learning rate was set to 10^{-3} and decreased following a plateau scheduler with a patience of ten epochs. Trainings were performed using the PyTorch library on an NVIDIA V100 GPU with 16 GB of memory.

2) Inference: Inference on the in silico dataset was executed using the ninefold cross-validation procedure, thus obtaining a prediction for each sample of the dataset. Ensemble inference was carried out for both the in vitro and in vivo datasets, by calculating the median value predicted for each pixel by the nine models of the cross-validation.

C. Metrics

To gauge the accuracy of Doppler velocity estimation, we calculated the root-mean-squared error (RMSE) between the ground-truth and predicted maps and examined its mean and standard deviation. Note that for the in vivo experiments, we refer to this metric as the root-mean-squared difference (RMSD) as we only had access to an imperfect reference, not the ground truth.

IV. RESULTS

A. In Silico Results

The performance of our methods in Doppler velocity estimation, when trained on Train-Set-1 and using a packet size $n = 2$, is presented in Table IV. For all methods, the metrics were computed exclusively within the left ventricle.

All deep learning approaches outperformed the autocorrelator. Real U-Net and Complex U-Net provided comparable

TABLE IV
DOPPLER VELOCITY ESTIMATION RESULTS OF THE BASELINE METHOD AND THE THREE PROPOSED DEEP LEARNING METHODS, TRAINED ON TRAIN-SET-1 AND EVALUATED ON ORIGINAL-SET, ALIASED-SET, AND ALIASED PIXELS OF ALIASED-SET. THE RESULTS CORRESPOND TO THE MEAN ± STANDARD DEVIATION OF THE RMSE (cm/s)

Method	Original-Set	Aliased-Set	Aliased pixels
Autocorrelator	3.7 ± 2.6	16 ± 9.8	57 ± 23
Real U-Net	1.8 ± 0.9	5.4 ± 5.3	14 ± 14
Complex U-Net	1.8 ± 0.8	$\mathbf{5.3 \pm 5.1}$	$\mathbf{14 \pm 13}$
ConvNeXt	$\mathbf{1.7 \pm 1.1}$	6.1 ± 6.3	16 ± 15

TABLE V
DOPPLER VELOCITY ESTIMATION RESULTS OF THE BASELINE METHOD AND THE THREE PROPOSED DEEP LEARNING METHODS TRAINED ON TRAIN-SET-2 AND EVALUATED ON ORIGINAL-SET, ALIASED-SET, AND ALIASED PIXELS OF ALIASED-SET. THE RESULTS REPRESENT THE MEAN ± STANDARD DEVIATION OF THE RMSE (cm/s)

Method	Original-Set	Aliased-Set	Aliased pixels
Autocorrelator	3.7 ± 2.6	16 ± 9.8	57 ± 23
Real U-Net	2.0 ± 0.8	2.4 ± 1.1	3.7 ± 3.1
Complex U-Net	2.0 ± 0.8	2.5 ± 1.3	3.7 ± 3.0
ConvNeXt	$\mathbf{1.8 \pm 0.7}$	$\mathbf{2.1 \pm 1.0}$	$\mathbf{3.3 \pm 2.4}$

results, with ConvNeXt slightly outperforming both. In addition, the variability of the three models was significantly lower than that of the autocorrelator.

We examined the samples with higher errors for both the autocorrelator and the deep learning models. For the former, large estimation errors mainly occurred in two situations: noisy signals and in the presence of aliasing. For the deep learning models, large errors were observed in two specific settings: in a few pixels near the valves and the presence of aliasing across a large region of the ventricle. To mitigate the effects of aliasing, we trained our models on Train-Set-2. The results of the estimation are presented in Table V.

We noted that when testing on Original-Set, we maintained a consistent performance compared to the models trained on Train-Set-1. Moreover, when testing on Aliased-Set, we observed a substantial reduction in error compared to the autocorrelator method, indicating that the models are capable to alleviate aliasing to some extent.

Illustrative examples are shown in Fig. 6. We report predictions obtained with ConvNeXt (the best-performing model) and Complex U-Net. The Real U-Net provided results similar to the Complex U-Net. As explained above, our models trained with both Train-Set-1 and Train-Set-2 achieved accurate estimates in the absence of aliasing (Samples A and B). In such cases, we observed that the autocorrelator was more susceptible to noise. When encountering small aliased regions (Sample C), the models trained with Train-Set-1 either effectively removed aliasing (see the Complex U-Net estimation) or tended to smooth the region, leading to velocity underestimation (see the right part of the ConvNeXt estimation). When trained with Train-Set-2, the estimation was closer to the ground truth (see the ConvNeXt estimation). In the case of larger aliased regions as in Sample D, the models trained on Train-Set-1 systematically smoothed the aliased region and hence underestimated velocities, whereas when trained with Train-Set-2, they yielded results closely aligned with the ground truth.

TABLE VI
DOPPLER VELOCITY ESTIMATION RESULTS OF THE BASELINE METHOD AND THE THREE PROPOSED DEEP LEARNING METHODS, EVALUATED ON STANDARD VOLUNTEER AND ALIASED VOLUNTEER SAMPLES. THE RESULTS REPRESENT THE MEAN ± STANDARD DEVIATION OF THE RMSD (cm/s)

Method	*Standard* samples	Aliased samples
Autocorrelator	$\mathbf{5.5 \pm 4.5}$	17 ± 8.8
Real U-Net	7.7 ± 5.4	9.8 ± 4.3
Complex U-Net	7.3 ± 4.9	11 ± 4.4
ConvNeXt	7.3 ± 4.6	$\mathbf{6.9 \pm 3.6}$

B. In Vitro Validation

Fig. 7 shows the RMSE obtained with the baseline and deep learning-based methods for all 12 in vitro experiments, using the same packet size $n = 2$. All models were trained on Train-Set-2. The RMSE is depicted as a function of the phantom outermost speeds. The Nyquist velocity v_N was kept unchanged across all experiments. We observed that our methods gave state-of-the-art results for $v < v_N$ and showed improved performance when $v > v_N$. Real U-Net and Complex U-Net gave comparable results, while ConvNeXt slightly underperformed them in terms of aliasing reduction for high rotation speeds.

Estimation maps for two of the experiments are presented in Fig. 8. In the first experiment, where the maximum velocity was below the Nyquist limit, we observed estimations closely aligned with the ground truth for all our networks. Moreover, our estimates were less noisy compared to those of the autocorrelator. In the second experiment, where the maximum velocity exceeded the Nyquist limit, the autocorrelator estimation clearly revealed the aliased regions. In contrast, our methods significantly reduced the aliased patterns. However, ConvNeXt displayed small noisy artifacts in the aliased regions.

C. In Vivo Validation

Estimation results on the in vivo sequences are reported in Table VI. We summarized results on 50 *standard* volunteer samples and 20 aliased samples, and the reference used for all RMSD calculations was the autocorrelation result with $n = 8$. All models were trained on Train-Set-2.

For *standard* samples, the autocorrelator with $n = 2$ produced the results closest to the reference. Fig. 9 shows the illustrative examples for Complex U-Net and ConvNeXt on one sample of each of the six subjects. From the samples of Subjects 1 and 4, all three methods displayed a similar estimation to the reference, but we observed that the two deep learning methods behaved as denoisers, unlike the autocorrelator with $n = 2$. Moreover, in the sample from Subject 3, the two deep learning estimations were also smoother than the reference obtained with $n = 8$.

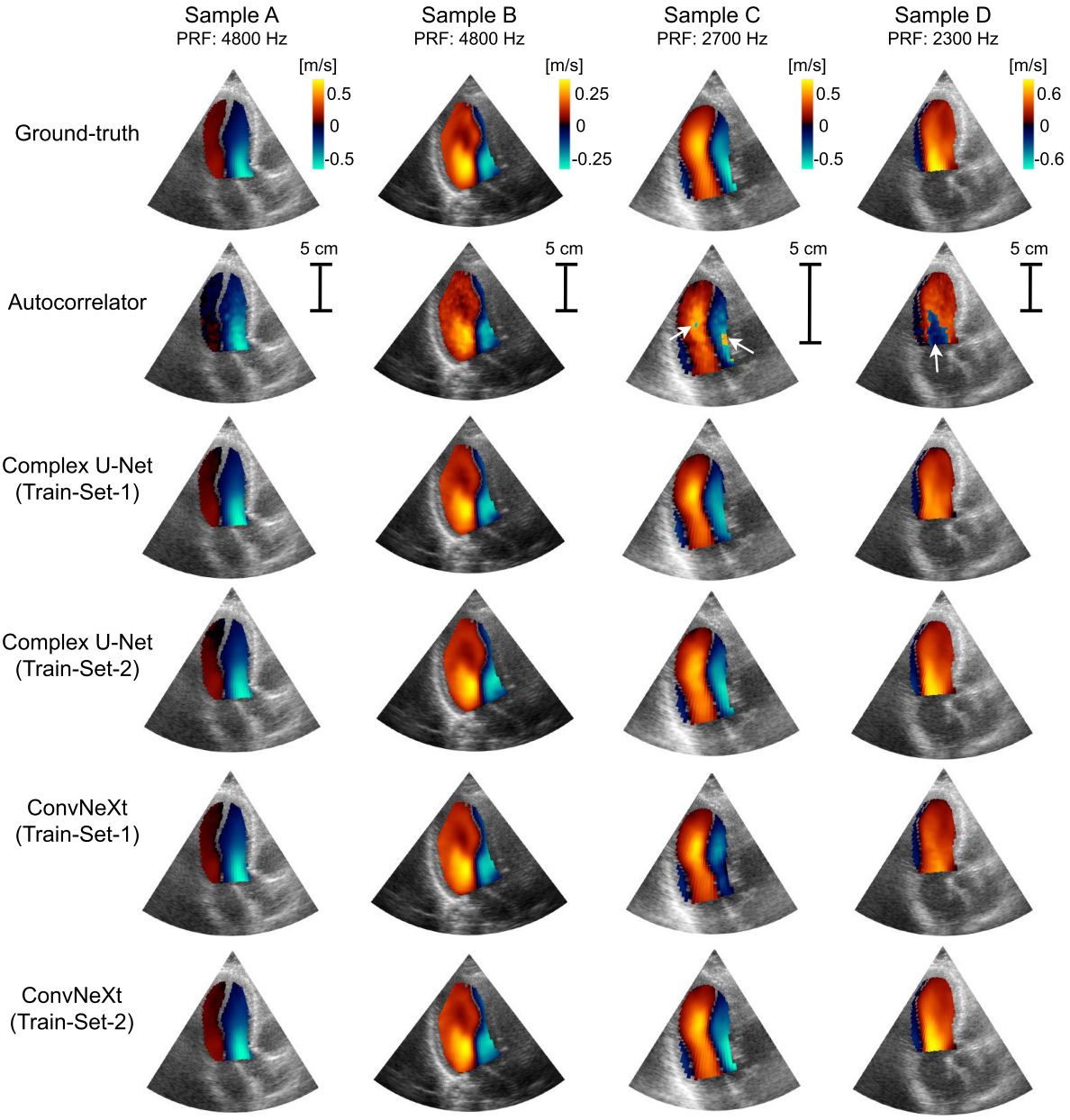


Fig. 6. Doppler velocity estimations for four in silico samples. B-mode and color Doppler images were simulated from the data of four different patients. Samples A and B belonged to Original-Set and had velocities below the Nyquist limit. Samples C and D belonged to Aliased-Set and had velocities above the Nyquist limit. Aliased regions are indicated with white arrows.

For aliased samples, as anticipated, the autocorrelator exhibited a substantial deviation from the reference due to aliasing, as reported in Table VI. Among the deep learning methods, ConvNeXt gave results significantly closer to the reference compared to the Real U-Net and the Complex U-Net. In Fig. 9, the sample from Subject 2 exemplifies a scenario with induced aliasing. The ConvNeXt reduced aliasing and produced an overall estimation very close to the reference, while the Complex U-Net yielded relatively large errors in the apical region and near the valve. In the sample from Subject 5, both Complex U-Net and ConvNeXt successfully reduced the induced aliasing. Finally, the sample from Subject 6 displays a sample with real aliasing during filling. Both the Complex U-Net and the ConvNeXt successfully reduced aliasing.

D. Computational Time

Table VII reports the sizes of the models and their training and inference times. With the same parameter configuration, Complex U-Net required significantly more training and inference time than Real U-Net. In contrast, ConvNeXt offered the most advantageous tradeoff between the number of parameters and the training and inference times. The inference times of the three models were sufficiently low to support real-time imaging.

V. Discussion

A. Doppler Velocity Estimation With Deep Learning

We achieved overall reliable real-time Doppler velocity estimates from a reduced packet size $n = 2$ with all proposed

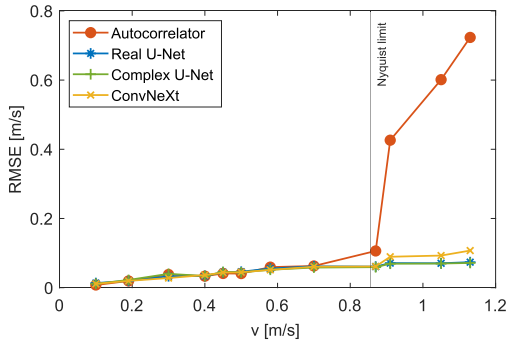


Fig. 7. Evolution of the RMSE with respect to the outermost speed of the phantom, obtained by the autocorrelator and deep learning methods. The RMSE is computed in the phantom area only. The Nyquist limit is indicated with a gray line.

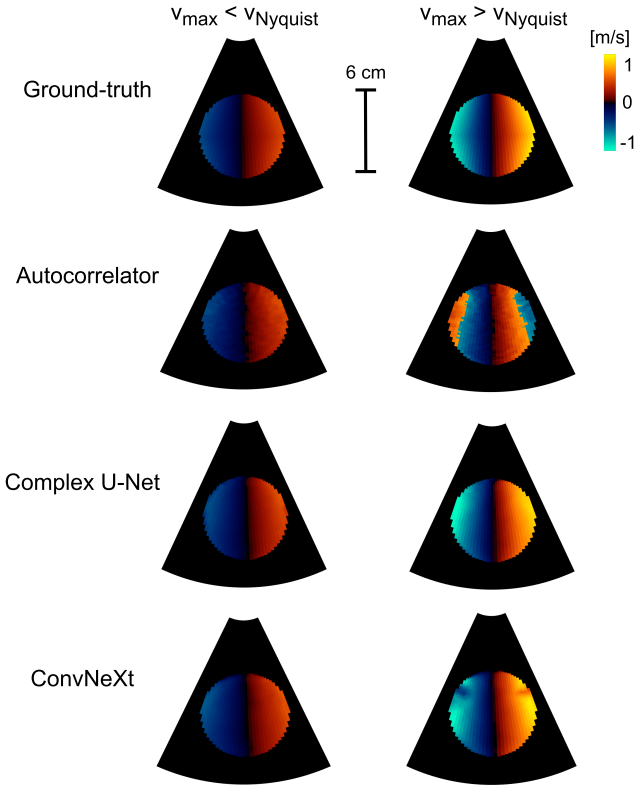


Fig. 8. Doppler velocity estimation results for two in vitro experiments with different rotating velocities. In the first one, $v_{max} = 0.45$ m/s. In the second one, $v_{max} = 1.15$ m/s. The Doppler fields are masked to show the phantom region only.

deep learning models. This held true for alias-free samples across in silico, in vivo, and in vitro data (Tables IV and VI and Fig. 7). When working with aliased sequences, we noticed that the models trained on alias-free data (Train-Set-1) tended to smooth the aliased regions (Sample D of Fig. 6), rendering postprocessing dealiasing methods less suitable. Augmenting the training dataset with aliased simulations notably improved the model's ability to mitigate aliasing in both in silico and in vitro data (Sample D of Fig. 6 and the second experiment of Fig. 8). For in vivo data, only ConvNeXt reliably reduced aliasing in all samples (Table VI and Fig. 9).

Regarding the autocorrelation method, our analysis revealed that our deep learning models yielded comparable quantitative results on in vitro and in vivo validations (Fig. 7 and Table VI, respectively). We also observed that this method remained stable when the packet size was reduced to $n = 2$, albeit producing some level of noise on the velocity maps (Subjects 1 and 4 in Fig. 9). However, in some cases, the estimation displayed a noticeable level of noise (Subject 3 in Fig. 9). In contrast, our method appeared to be more robust to noise (as evident in Subjects 1, 3, and 4 in Fig. 9).

TABLE VII
COMPARISON OF THE COMPLEXITY OF THE AUTOCORRELATOR AND THE THREE DEEP LEARNING MODELS. THE TRAINING TIMES ARE REPORTED FOR THE TRAINING OF ONE FOLD, ASSUMING THAT ALL FOLDS CAN BE TRAINED IN PARALLEL. THE INFERENCE TIMES ARE REPORTED FOR THE INFERENCE OF IN VIVO FRAMES. TRAINING AND INFERENCE WERE BOTH CONDUCTED ON GPU FOR DEEP LEARNING MODELS AND ON CPU FOR THE AUTOCORRELATOR

Model	Parameters	Training time	Inference time
Autocorrelator	x	x	$\sim$ 1.6 ms
Real U-Net	1.5 M	$\sim$ 0.5 h	$\sim$ 2.5 ms
Complex U-Net	1.5 M	$\sim$ 5 h	$\sim$ 18 ms
ConvNeXt	4.7 M	$\sim$ 1 h	$\sim$ 5.3 ms

A critical concern with deep learning methods is their capacity to generalize. Models frequently fail to extrapolate when the distribution shift between training and validation datasets is too large. This issue is particularly pronounced in medical imaging, where distribution shifts can arise from various factors, such as differences in data acquisition or preprocessing, and training with unrealistic simulations. In our study, we obtained encouraging results regarding the generalization capability of our models, despite being trained exclusively on in silico samples. Notably, our models demonstrated high-quality results on unseen nonflow in vitro data and on in vivo data. We only observed large errors in a limited number of cases (as exemplified by Subject 2 in Fig. 9).

This highlights the effectiveness of our simulation pipeline in generating color Doppler samples to build supervised training datasets when no ground truth is available. Our pipeline enables users to expand the range of clinical scenarios, including different echocardiographic views, transmission types (e.g., focused and diverging), and transducer types. It uses patient-derived Doppler echocardiographic data to model cardiac tissue and blood backscatterers. Although our focus was on three-chamber views, other views could thus be explored. In addition, the SIMUS simulator provides the flexibility to specify various probe and transmission parameters, allowing for the simulation of ultrasound signals with nonfocused waves. This capability facilitates the generation of a vast pool of training data, enabling the creation of more generalized models. It is also worth noting that the favorable results obtained in nonflow data during the in vitro disk experiments suggest the potential for directly applying our models to other modalities involving Doppler velocity estimation, such as tissue Doppler.

The capacity for generalization is also contingent upon the size of the training dataset. Data augmentation is a crucial strategy for enhancing generalization ability, as it reduces the risk of overfitting the training data by increasing the size and diversity of the training dataset. While typical augmentation

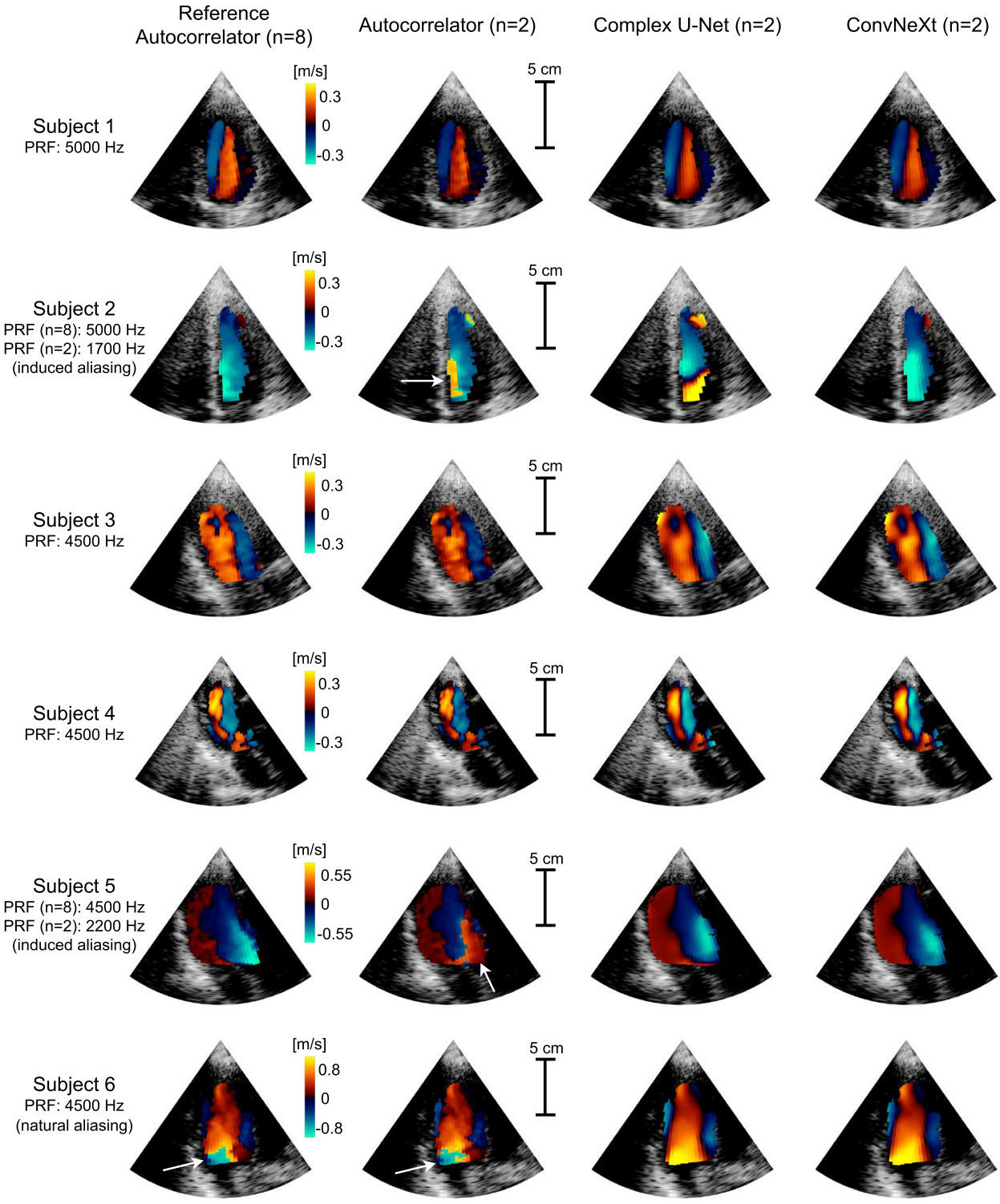


Fig. 9. Doppler velocity estimation results on an in vivo sample of six different subjects, using the autocorrelator with a packet size of 8 as a reference. In Subjects 2 and 5, samples with $n = 2$ were constructed with a decreased PRF to induce aliasing. The sample from Subject 6 naturally contained aliasing. Aliased regions are indicated with white arrows.

techniques, such as rotation or cropping, can be readily applied to images, they can also result in significant distortion of data such as I/Q signals. Therefore, we have developed tailored augmentation procedures to increase the size of the dataset while maintaining realistic signals. We observed experimentally that increasing the number of augmented samples was linked to enhanced generalization performances.

Finally, this work could offer practical advantages for focused color Doppler echocardiography. For instance, assuming a typical packet size of around $n = 8$, reducing the packet

size to $n = 2$ could result in a fourfold increase in frame rate, yielding approximately 60 frames/cardiac cycle. Such an enhancement could ease quantitative assessment of blood velocity throughout the cardiac cycle, potentially simplifying the application of techniques such as intracardiac vector flow mapping.

B. Model Comparison

Two data representation methods were evaluated: real-valued and complex-valued. In addition, two types of architectures were benchmarked: U-Net and ConvNeXt.

1) Real-Valued Versus Complex-Valued Representations: We anticipated that Complex U-Net would outperform Real U-Net, in line with findings from our previous studies [28]. However, Real U-Net and Complex U-Net yielded comparable results. One hypothesis is that while the Doppler velocity estimation problem involves complex-valued inputs, the output data are real-valued. Therefore, a transition from complex-valued to real-valued representations may be necessary, and a complex-valued architecture may not fully leverage its capabilities in this specific scenario. We also noted that the use of complex-valued representations significantly increased computational time during both training and inference due to the higher number of operations associated with complex-valued convolutions.

2) U-Net Versus ConvNeXt: In this study, we analyzed ConvNeXt, a state-of-the-art deep learning architecture built on enhanced convolutional layers mirroring Transformer blocks for improved performance. We found similar performance from U-Net and ConvNeXt on in silico and in vitro data. Although ConvNeXt exhibited imperfect results on highly aliased in vitro samples, it demonstrated greater robustness when applied to in vivo aliased samples. This performance indicates the greater ability of ConvNeXt to learn intracardiac flow geometry from in silico data. One of the reasons may be the larger number of training parameters of the ConvNeXt, allowed by the use of the efficient depthwise separable convolutions, which require less computational operations than standard convolutions. Hence, ConvNeXt may generalize better to intracardiac in vivo data than to nonflow phantom data, which is highly out-of-distribution compared to training data. Moreover, despite including more parameters, ConvNeXt provided fast training and inference times.

C. Limitations and Perspectives

Despite the positive findings presented in this article, we acknowledge two main limitations of this work. First, the in vivo results must be regarded as preliminary as the absence of ground truth made it difficult to draw firm conclusions on estimation quality. Second, we did not account for clutter influence and assumed that clutter filtering could be performed effectively on a packet size of two. A challenging perspective is to develop a clutter filtering method specifically designed for reduced packet sizes. We believe that exploring deep learning techniques could be a promising avenue for addressing this task.

Another perspective is to incorporate temporal information [18]. At this stage, our simulations were based on patient data from an echocardiographic laboratory and thus lacked sufficient temporal resolution to assume continuity between consecutive frames. Realistic color flow imaging simulations based on a numerical heart, such as those employed previously in our research [17], [29], would provide the necessary frame rate to leverage temporal information effectively and improve the robustness of our models.

Finally, our models are expected to generalize to pathological flows, especially if nonstandard flows are present in the training dataset. However, due to anonymization, we cannot determine which or how many of the patients used to generate in silico data have a cardiac condition. Also, the in vivo validation was exclusively performed on healthy volunteers. Therefore, a validation study on a cohort, including a variety of pathological cases, could be envisaged for future work. In particular, it would allow us to assess the robustness of ConvNeXt when encountering highly abnormal flows, as it showed suboptimal performance when processing out-of-distribution in vitro data.

VI. Conclusion

We explored deep learning models for Doppler velocity estimation from clutter-filtered I/Q signals with a reduced packet size of $n = 2$. Using a color Doppler simulation pipeline, we built a supervised training dataset and applied custom data augmentation, including artificially adding aliased samples. Our method exhibited superior performance on in silico data compared to the autocorrelation method and achieved proficient results on in vitro and in vivo data. We demonstrated the effectiveness of our approach in reducing aliasing and handling noise. Three state-of-the-art deep learning strategies were adapted and evaluated, with ConvNeXt showing the best generalization performance on in vivo samples. Overall, our study highlights the potential of deep learning for quantitative color Doppler analysis with minimal data.

Acknowledgment

The authors would like to thank Adeline Bernard for the phantom design, Nin Ghigo for their help with in vivo acquisitions, and Prof. Olivier Bernard, Thierry Judge, and Florian Vixège for their help in the development of the simulation pipeline. This work was performed using HPC resources from GENCI–IDRIS (Grant 2022-AD011013679). The in vitro material is based upon work done on the ISO 9001:2015 PILoT facility. The RF Verasonics generator was cofounded by the FEDER program, Saint-Etienne Metropole (SME), and Conseil General de la Loire (CG42) within the framework of the SonoCardio-Protection Project led by Prof. Pierre Croisille.

References

[1] C. Mitchell et al., "Guidelines for performing a comprehensive transthoracic echocardiographic examination in adults: Recommendations from the American Society of Echocardiography," *J. Amer. Soc. Echocardiogr.*, vol. 32, no. 1, pp. 1–64, Jan. 2019.

[2] L. Jerome Krovetz and S. D. Goldbloom, "Frequency content of intravascular and intracardiac pressures and their time derivatives," *IEEE Trans. Biomed. Eng.*, vol. BME-21, no. 6, pp. 498–501, Nov. 1974.

[3] J. Faurie, M. Baudet, J. Porée, G. Cloutier, F. Tournoux, and D. Garcia, "Coupling myocardium and vortex dynamics in diverging-wave echocardiography," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 66, no. 3, pp. 425–432, Mar. 2019.

[4] A. C. H. Yu, K. W. Johnston, and R. S. C. Cobbold, "Frequency-based signal processing for ultrasound color flow imaging," *Canadian Acoust.*, vol. 35, no. 2, pp. 11–23, 2007.

[5] T. Loupas, J. T. Powers, and R. W. Gill, "An axial velocity estimator for ultrasound blood flow imaging, based on a full evaluation of the Doppler equation by means of a two-dimensional autocorrelation approach," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 42, no. 4, pp. 672–688, Jul. 1995.

[6] D. Posada et al., "Staggered multiple-PRF ultrafast color Doppler," *IEEE Trans. Med. Imag.*, vol. 35, no. 6, pp. 1510–1521, Jun. 2016.

[7] S. Muth, S. Dort, I. A. Sebag, M.-J. Blais, and D. Garcia, "Unsupervised dealiasing and denoising of color-Doppler data," *Med. Image Anal.*, vol. 15, no. 4, pp. 577–588, Aug. 2011.

[8] B. Luijten, N. Chennakeshava, Y. C. Eldar, M. Mischi, and R. J. G. van Sloun, "Ultrasound signal processing: From models to deep learning," *Ultrasound Med. Biol.*, vol. 49, no. 3, pp. 677–698, Mar. 2023.

[9] R. J. Van Sloun, H. Belt, K. Janse, and M. Mischi, "Learning Doppler with deep neural networks and its application to intra-cardiac echography," in *Proc. IEEE Int. Ultrason. Symp. (IUS)*, 2018, pp. 1–4.

[10] I. Z. Apostolakis et al., "Systems and methods for generating color Doppler images from short and undersampled ensembles," U.S. Patent 2 02 288 73A1, Jul. 20, 2023.

[11] B. He, J. Lei, X. Lang, Z. Li, W. Cui, and Y. Zhang, "Ultra-fast ultrasound blood flow velocimetry for carotid artery with deep learning," *Artif. Intell. Med.*, vol. 144, Oct. 2023, Art. no. 102664.

[12] O. Solomon et al., "Deep unfolded robust PCA with application to clutter suppression in ultrasound," *IEEE Trans. Med. Imag.*, vol. 39, no. 4, pp. 1051–1063, Apr. 2020.

[13] K. G. Brown, D. Ghosh, and K. Hoyt, "Deep learning of spatiotemporal filtering for fast super-resolution ultrasound imaging," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 67, no. 9, pp. 1820–1829, Sep. 2020.

[14] H. Nahas, J. S. Au, T. Ishii, B. Y. S. Yiu, A. J. Y. Chee, and A. C. H. Yu, "A deep learning approach to resolve aliasing artifacts in ultrasound color flow imaging," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 67, no. 12, pp. 2615–2628, Dec. 2020.

[15] H. Nahas, B. Y. S. Yiu, A. J. Y. Chee, J. S. Au, and A. C. H. Yu, "Deep-learning-assisted and GPU-accelerated vector Doppler imaging with aliasing-resistant velocity estimation," *Ultrasonics*, vol. 134, Sep. 2023, Art. no. 107050.

[16] H. J. Ling, O. Bernard, N. Ducros, and D. Garcia, "Phase unwrapping of color Doppler echocardiography using deep learning," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 70, no. 8, pp. 810–820, Aug. 2023.

[17] Y. Sun et al., "A pipeline for the generation of synthetic cardiac color Doppler," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 69, no. 3, pp. 932–941, Mar. 2022.

[18] H. J. Ling, N. Painchaud, P.-Y. Courand, P.-M. Jodoin, D. Garcia, and O. Bernard, "Extraction of volumetric indices from echocardiography: Which deep learning solution for clinical use?" in *Proc. FIMH*, 2023, pp. 54–245.

[19] F. Vixège et al., "Physics-constrained intraventricular vector flow mapping by color Doppler," *Phys. Med. Biol.*, vol. 66, no. 24, Dec. 2021, Art. no. 245019.

[20] D. Garcia, "SIMUS: An open-source simulator for medical ultrasound imaging. Part I: Theory & examples," *Comput. Methods Programs Biomed.*, vol. 218, May 2022, Art. no. 106726.

[21] A. Cigier, F. Varray, and D. Garcia, "SIMUS: An open-source simulator for medical ultrasound imaging. Part II: Comparison with four simulators," *Comput. Methods Programs Biomed.*, vol. 220, Jun. 2022, Art. no. 106774.

[22] D. Garcia, "Make the most of MUST, an open-source MATLAB UltraSound toolbox," in *Proc. IEEE Int. Ultrason. Symp. (IUS)*, Xi'an, China, Sep. 2021, pp. 1–4.

[23] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Med. Image Comput. Comput.-Assist. Intervent.*, 2015, pp. 234–241.

[24] C. Trabelsi et al., "Deep complex networks," in *Proc. ICLR*, 2018, pp. 1–19.

[25] M. W. Matthés, Y. Bromberg, J. de Rosny, and S. M. Popoff, "Learning and avoiding disorder in multimode fibers," *Phys. Rev. X*, vol. 11, Apr. 2021, Art. no. 021060.

[26] Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, "A ConvNet for the 2020s," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 11966–11976.

[27] F. Mehregan et al., "Doppler vortography: A color Doppler approach to quantification of intraventricular blood flow vortices," *Ultrasound Med. Biol.*, vol. 40, no. 1, pp. 210–221, Jan. 2014.

[28] 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 Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 69, no. 2, pp. 592–603, Feb. 2022.

[29] C. Chnafa, S. Mendez, R. Moreno, and F. Nicoud, "Using image-based CFD to investigate the intracardiac turbulence," in *Modeling the Heart and the Circulatory System*, vol. 14, A. Quarteroni, Ed., Cham, Switzerland: Springer, 2015, pp. 97–117.