



HAL
open science

Complex Convolutional Neural Networks for Fast Diverging Wave Imaging

Jingfeng Lu, Fabien Millioz, Damien Garcia, Sébastien Salles, Dong Ye, Denis Friboulet

► **To cite this version:**

Jingfeng Lu, Fabien Millioz, Damien Garcia, Sébastien Salles, Dong Ye, et al.. Complex Convolutional Neural Networks for Fast Diverging Wave Imaging. 2020 IEEE International Ultrasonics Symposium (IUS 2020), IEEE, Sep 2020, Las Vegas, NV, United States. 10.1109/IUS46767.2020.9251325 . hal-02353669

HAL Id: hal-02353669

<https://hal.science/hal-02353669>

Submitted on 23 Jun 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Complex Convolutional Neural Networks for Fast Diverging Wave Imaging

Jingfeng Lu^{*†}, Fabien Millioz[†], Damien Garcia[†], Sébastien Salles[†], Dong Ye^{*}, and Denis Friboulet[†]
 Email: jingfeng.lu@hit.edu.cn, {fabien.millioz, damien.garcia, sebastien.salles}@creatis.insa-lyon.fr,
 yedong@hit.edu.cn, denis.friboulet@creatis.insa-lyon.fr

^{*}Metislab, School of Instrumentation Science and Engineering, Harbin Institute of Technology, Harbin, China

[†]University of Lyon, CREATIS, CNRS UMR 5220, Inserm U1044, INSA-Lyon, University of Lyon 1, Lyon, France

Abstract—Single diverging wave (DW) imaging produces ultrasound (US) images at high frame rate (ultrafast) but of low quality. Conventional high-quality DW imaging relies on the coherent compounding of multiple consecutive steered emissions, which in turn reduces the gain in frame rate. Reconstructing high-quality US images for ultrafast imaging using deep learning techniques has recently raised a growing interest in the US community. We recently described a convolutional neural network (CNN) architecture called ID-Net, which exploited an inception layer devoted to the reconstruction of DW ultrasound images using radio frequency (RF) data. We derive in this work the complex equivalent of this network, i.e., the complex inception for DW network (CID-Net), operating on in-phase/quadrature (I/Q) data. We experimentally demonstrate that the CID-Net yields the same image quality as that obtained from the RF-trained CNN, i.e., using only three I/Q images, the CID-Net yields high-quality images competing with those obtained by coherently compounding 31 RF images.

Index Terms—Deep learning, complex convolutional neural networks (CCNNs), ultrasound imaging, diverging wave, image reconstruction.

I. INTRODUCTION

Ultrasound (US) images obtained from single diverging wave (DW) emissions exhibit low image quality. A conventional reconstruction approach consists in coherently compounding series of ultrasound signals from steered DW transmissions [1], at the expense of frame rate, data volume, and computation time. Recently, there has been a growing interest in applying deep learning techniques to improve US imaging. Most of the existing studies operate on radio frequency (RF) signals [2], while modern ultrasound systems typically sample the complex in-phase/quadrature (I/Q) baseband signals instead of RF signals. Thus in this work, inspired by the study of Trabelsi et al. [3], we propose to extend the deep learning-based US imaging to the complex domain using complex convolutional neural networks (CCNNs).

We present the complex inception for DW network (CID-Net) constructed with the complex building components introduced in [3], for high-quality DW image reconstruction from I/Q data. CID-Net is derived from our prior work [4], [5], inception for DW Network (ID-Net), which has demonstrated the ability to reconstruct high-quality DW images from RF data. We provide experimental evidence that the CID-Net yields the same image quality as that obtained from the

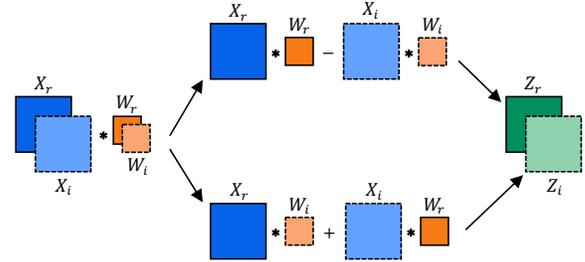


Fig. 1. Block diagram of complex convolution. The orange, blue, and green blocks denote the convolution kernel W , input data X , and output data Z . Both the data and kernel are formed as the concatenation of two real-valued matrixes, each representing the real (denoted as solid blocks) and imaginary (denoted as dotted blocks) parts.

RF-trained CNN, and an image quality competing with the coherent compounding of 31 DWs.

II. METHODS

The DW image reconstruction is modeled as an image input-output problem where the objective is to estimate a high-quality image \hat{Y} using I/Q data X obtained from m DW transmissions. We propose to use the CID-Net with trainable complex-valued parameters Θ to seek for the optimal compounding operation $\hat{Y} = F(X; \Theta)$, with respect to the high-quality target images Y obtained from the coherent compounding of n ($n \gg m$) DWs.

A. Complex Convolution

As depicted in Fig. 1, we used real-valued matrixes to represent the real and imaginary components of complex data, and performed complex convolution using real-valued arithmetic. We consider a complex-valued data matrix $X = X_r + jX_i$, where $j = \sqrt{-1}$ is the imaginary unit, while $X_r = \text{Re}(X)$ and $X_i = \text{Im}(X)$ are the real and imaginary components of X respectively. Likewise, we represent the complex-valued weight of a convolution kernel as $W = W_r + jW_i$. We convolve W with X as

$$Z = W * X = (W_r + jW_i) * (X_r + jX_i). \quad (1)$$

Considering the distributive property of convolution, the complex convolution is represented with four real convolutions as

$$Z = (W_r * X_r - W_i * X_i) + j(W_r * X_i + W_i * X_r). \quad (2)$$

TABLE I
ARCHITECTURE OF PROPOSED NETWORK

layer type	feature size channel \times height \times width	kernel size height \times width	nb. of kernels
inputs	$m \times h \times w$	-	-
convolution	$32 \times h \times w$	9×3	128
convolution	$16 \times h \times w$	17×5	64
convolution	$8 \times h \times w$	33×9	32
inception	$4 \times h \times w$	41×11	4
		49×13	4
		57×15	4
convolution	$1 \times h \times w$	65×17	4
		1×1	4

Thus the mathematical relation between real and imaginary components is fully reflected in this representation.

B. Network Architecture

We constructed the CID-Net using the complex convolution blocks described in Section II-A. The CID-Net maintained the architecture of ID-Net, which was devoted to the reconstruction of DW ultrasound images from RF data. The architecture was a fully convolutional network composed of five complex convolution layers with maxout activation units. Particularly, the second last layer was an inception layer consisting of multi-size convolution kernels. As demonstrated in [4], the inception layer used in conjunction with maxout activation units allowed features from multiple receptive field sizes to be captured, which contributed to dealing with the specific geometry of DW. A more detailed specification of the architecture is provided in Table I.

III. EXPERIMENTS

A. Dataset Acquisition

We performed steered DW acquisitions using a Verasonics system research scanner (Vantage 256) equipped with an ATL P4-2 probe (bandwidth: 2-4 MHz, center frequency: 3 MHz). The acquisition planes were obtained by continuously moving the probe on the surface of target areas, at an imaging rate of 50 frames/s and a packet size of 250 frames. Each acquisition plane was obtained using 31 steered DWs spanning $\pm 30^\circ$ in 2° steps. The received raw data were I/Q demodulated, then downsampled by a factor 2 and beamformed using a delay-and-sum (DAS) to produce the beamformed I/Q data. The input X of CID-Net were made up of $m = 3$ I/Q images corresponding to steering angles (-20° , 0° , and 20°), while the reference images Y were the compounding of all $n = 31$ I/Q images. The data used in the experiments corresponds to 7500 (X , Y) samples. Specifically, 1500 *in-vivo* samples were acquired from three healthy subjects (thigh muscle, finger phalanx, and liver regions), and 6000 *in-vitro* samples were acquired from two phantoms (Gammex, model 410SCG, and CIRS, model 054GS). From the 7500 (X , Y) samples, 5000 samples were used as the training set and 1250 samples were

used as an independent validation set. The remaining 1250 samples were used for testing.

B. Training Implementation

The training was implemented with Pytorch [6] library on an NVIDIA Tesla V100 GPU with 32 Gb of memory. Specifically, the network weights were initialized with the Xavier initializer [7]. Mean squared error (MSE) was used as the training loss and minimized using mini-batch gradient descent with the Adam optimizer [8]. The batch size was set to 16, and the initial learning rate was set to 1×10^{-4} . During the training process, the learning rate was halved if there had been no decrease in the validation loss for 20 epochs, and 40 epochs without validation loss reduction would end the training, resulting in training time of two days.

C. Evaluation Metrics

We used contrast ratio (CR), contrast-to-noise ratio (CNR), and lateral resolution (LR) to quantitatively evaluate the performance of the proposed network.

CR and CNR were used to measure the contrast between the object of interest and the surrounding background.

$$\text{CR} = -20 \log_{10} \frac{\mu_t}{\mu_b}, \quad (3)$$

$$\text{CNR} = 20 \log_{10} \frac{|\mu_t - \mu_b|}{\sqrt{\sigma_t^2 + \sigma_b^2}}, \quad (4)$$

where μ_t and μ_b (σ_t^2 and σ_b^2) denote the means (variances) of the intensity within the target region and the background. The CR and CNR were measured on two cysts (in the near field at 40-mm depth and the far field at 120-mm depth) of the images obtained from the Gammex phantom.

LR was used to assess the width of the point spread function from point target images. The full width at half maximum was used in this work to assess the LR. The LR was measured on 0.1-mm nylon monofilaments (in the near field at 20-mm and 40-mm depth, the middle field at 60-mm depth, and the far field at 80-mm, 90-mm, and 100-mm depth) of the images obtained from the CIRS phantom.

IV. RESULTS

Fig. 2 displays the B-mode images of the representative *in-vivo* samples, obtained from RF data using the compounding method and ID-Net, as well as the image obtained from I/Q data using the CID-Net. It can be observed that the ID-Net and CID-Net both produced better image quality than that of the compounding with the same three DWs, yielding images visually very close to the reference image.

We report in Table II the CR, CNR, and LR reached by the compounding method, CID-Net, and ID-Net. CID-Net and IDNet obtained approximately the same values in all evaluation metrics, while providing slightly lower values in CR and CNR, and lower values in LR than those associated to the reference.

TABLE II
EVALUATION METRICS OF COMPOUNDING METHOD, CID-NET, AND ID-NET [4].

model	CR [dB]		CNR [dB]		LR [mm]		
	near field	far field	near field	far field	near field	middle field	far field
compounding (3 DWs)	12.24	10.54	2.94	3.02	1.05	1.54	1.94
CID-Net (3 IQ images)	21.51	18.24	8.10	6.35	1.05	1.67	2.08
ID-Net (3 RF images)	21.47	18.48	8.11	6.32	1.06	1.67	2.04
compounding (31 DWs)	21.74	18.85	8.20	6.45	1.13	1.76	2.20

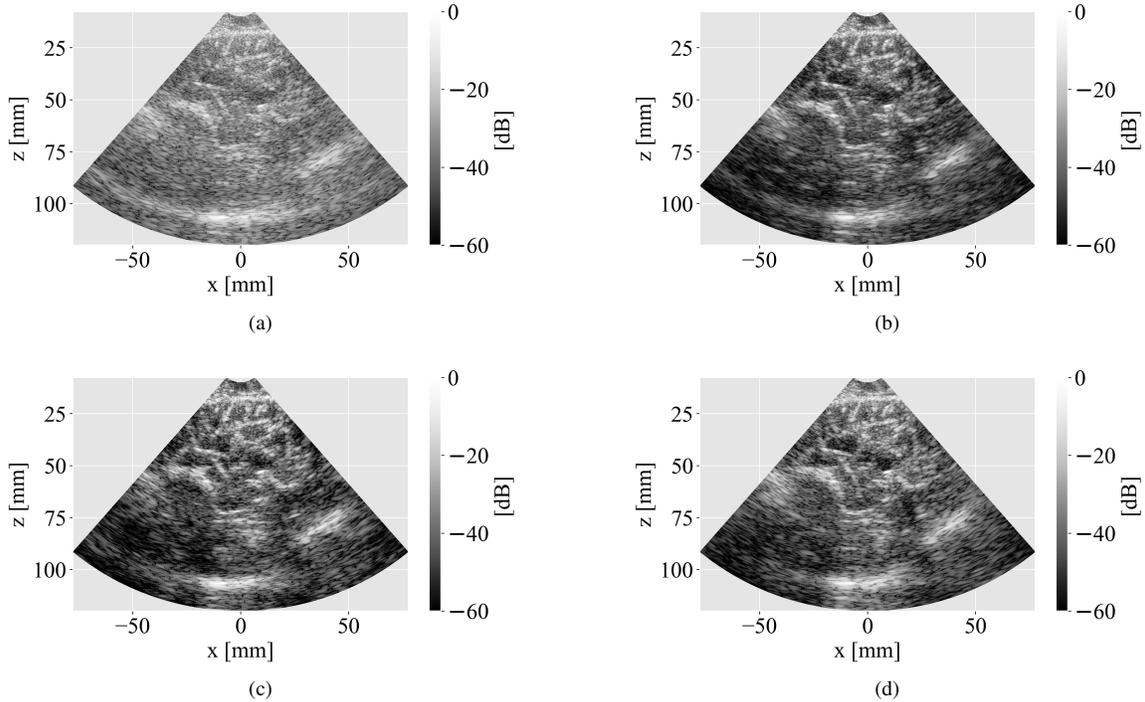


Fig. 2. B-mode images obtained with (a) compounding of 3 DWs, (b) CID-Net, (c) ID-Net [4], and (d) compounding of 31 DWs (reference), acquired from *in-vivo* scans (thigh muscle).

V. CONCLUSION

We proposed a complex convolutional neural network (CCNN), CID-Net, for fast and high-quality diverging wave (DW) imaging. The network was trained to learn a compounding operator for reconstructing high-quality images from I/Q data obtained with a small number of DW transmissions. Experimental results showed that the proposed CID-Net yielded the same image quality as the real equivalent CNN trained with RF data, and an image quality competing with the coherent compounding of 31 DWs. The proposed work will provide motivation to explore CCNN-based approaches for US imaging applications.

REFERENCES

- [1] M. Tanter and M. Fink, "Ultrafast imaging in biomedical ultrasound," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 61, no. 1, pp. 102–119, 2014.
- [2] 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, 2017.
- [3] C. Trabelsi, O. Bilaniuk, Y. Zhang, D. Serdyuk, S. Subramanian, J. F. Santos, S. Mehri, N. Rostamzadeh, Y. Bengio, and C. J. Pal, "Deep complex networks," in *International Conference on Learning Representations*, 2018.
- [4] 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*, 2020.
- [5] 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, 2019, pp. 2341–2344.
- [6] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, "Automatic differentiation in pytorch," in *Advances in Neural Information Processing Systems Workshop*, 2017.
- [7] 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.
- [8] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proceeding of International Conference on Learning Represent*, 2015, pp. 1–41.