Memory-efficient neural network for non-linear ultrasound computed tomography reconstruction

Fan, Yuling; Wang, Hongjian; Gemmeke, Hartmut; Hopp, Torsten; Dongen, Koen Van; Hesser, Jurgen

DOI
10.1109/ISBI48211.2021.9434164

Publication date
2021

Document Version
Final published version

Published in
2021 IEEE 18th International Symposium on Biomedical Imaging, ISBI 2021

2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)
April 13-16, 2021, Nice, France

MEMORY-EFFICIENT NEURAL NETWORK FOR NON-LINEAR ULTRASOUND COMPUTED TOMOGRAPHY RECONSTRUCTION

Yuling Fan[1], Hongjian Wang[2], Hartmut Gemmeke[3], Torsten Hopp[3], Koen van Dongen[4], Jürgen Hesser[1]

1. Heidelberg University, Mannheim, Germany
2. Donghua University, Shanghai, China
3. Karlsruhe Institute of Technology, Karlsruhe, Germany
4. Delft University of Technology, Delft, Netherlands

2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI) | 978-1-6654-1246-9/20/$31.00 ©2021 IEEE | DOI: 10.1109/ISBI48211.2021.9434164

ABSTRACT

Deep neural networks have proven to excel classical medical image reconstruction techniques. Some networks are based on fully connected (FC) layers to achieve domain transformation such as from the data acquisition domain to the image domain. However, FC layers result in huge numbers of parameters which take a lot of GPU memory. Hence, they do not scale well, and the overall performance is limited. For ultrasound computed tomography (USCT) application, we propose a memory-efficient convolutional network that reconstructs images from the frequency domain to image domain with much less parameters compared with multilayer perceptron, by using data-driven learning. Extensive experiments demonstrate that our method achieves high reconstruction quality. It improves the structural similarity measure (SSIM) from 0.73 to 0.99 when compared with state-of-the-art reconstruction methods in this field while reduces 2/3 parameters when compared with deep neural network with FC layers to reconstruct images from frequency domain to image domain.

Index Terms— ultrasound computed tomography, memory efficiency, image reconstruction, deep learning

1. INTRODUCTION

USCT offers high potential for early breast cancer detection without the use of ionizing radiation. Image acquisition of USCT uses several thousand ultrasound transducer elements distributed over a ring or a hemisphere, surrounding the breast submerged in water [1]. Each transducer sequentially emits an unfocused ultrasound wave, whereby all other transducers receive the transmitted and reflected signals as A-scans. While traversing and interacting with breast tissue, amplitude and phase of the acoustic wave change. Repeated data acquisitions for all emitters result in several million A-scans, and deliver the necessary data for reconstruction. The reconstruction of USCT images is an ill-posed nonlinear inverse problem, and the reconstruction is very time-consuming, which limits clinical implementations.

USCT reconstruction is a time-consuming, highly non-linear inverse problem. Conventional reconstruction methods iteratively compute the expensive forward model and its gradient. Reconstruction speed depends strongly on the number of such iterations. It (so far) turned out to be hard for deep learning strategies to learn the non-linear transformation from the image acquisition domain to image domain without conventional reconstruction as input or FC layers [4] to obtain high-resolution reconstructed images.

In this paper, we demonstrate how to overcome this limitation by two measures: (1) the new network does not rely on conventional reconstruction for domain transformation, which reduces the computing time; (2) it can learn the partial inverse with much fewer parameters, which makes possible for large images, offering a powerful tool for solving reconstruction problems in clinically acceptable time.

2. FORWARD MODEL

For conventional iterative image reconstruction and model-based deep neural networks, it is necessary to compute the forward model. In USCT, the basis for forward model is the wave propagation of ultrasound which is mathematically described by the wave equation (e.g. in the frequency domain) for an inhomogeneous object [1]. We consider the Helmholtz equation for ultrasound propagation through an acoustic background medium including refraction, diffraction and multiple forward scattering as

$$\Delta p + k_0^2 (1+\eta)^2 p = 0 \qquad (1)$$

with the frequency-dependent pressure filed p. The acoustic medium is described by the background wave number $k_0 = w/c$ and the refractive index $1+\eta$, where $\omega = 2\pi f$ is the angular frequency for frequency f and speed of sound (SoS) c of the background medium, and η accounts for the deviation of the inhomogeneity from the background medium.

The goal of USCT image reconstruction is to reconstruct SoS and attenuation, which are incorporated in the complex variable η that describes the deviation of SoS and attenuation in the breast from the background medium of water.

2.1. Paraxial approximation forward model

The Helmholtz equation is highly computationally demanding and it is not possible to be solved within a reasonable time due to the used maximal frequencies of up to 3.5 MHz requiring approximately 10 discretization points per wavelength [1]. An acceleration strategy, commonly known from geophysics, is the paraxial approximation [1], which describes the ultrasound field as nearly planar waves in forward direction. According to the paraxial approximation, the frequency dependent pressure field p reads

$$p_{k+1} = e^{i\Delta z k_0 \eta_k} * \mathbf{F}^{-1}\left\{ e^{i\Delta z\sqrt{k_0^2-\xi^2}} * \mathbf{F}(p_k) \right\} \qquad (2)$$

the index k at p and η denotes the k$^{\text{th}}$ z slice, the spectral variable $\xi = \frac{2\pi}{\Delta x N x}\left[-\frac{Nx}{2}+1,\ldots,0,\ldots,\frac{Nx}{2}\right]^T \in \mathrm{R}^{Nx}$, and the 1D discrete Fourier transformations and the inverse discrete Fourier transformations are denoted by $\mathbf{F}$ and $\mathbf{F}^{-1}$.

3. METHOD

The efficiency of U-Net's with their U-shaped structure has been demonstrated in image denoising [3] and is attributed to its better GPU memory efficiency, large receptive field, and lower computational cost [6]. Inspired by it, six U-shaped structures are used to increase the overall size of the feature maps and replace the FC layers with less parameters while reconstructing images from frequency domain to image domain.

The architecture, shown in Figure 1, consists of two parts: *down-up sample block* (DUB) and *reconstruction block*. In the DUB (see Figure.1(b)), a convolutional layer with a 3 × 3 kernel and stride of 2 is followed by a rectified linear unit (ReLU) to achieve down-sampling. During down-sampling, the number of feature map is doubled. In the progress of up-sampling, a subpixel layer is used, since the number of features is reduced by a quarter through the sub-pixel layer. The number of feature maps can be increased through a 1×1 convolution layer before the sub-pixel layer to maintain information. The DUBs are inserted between two modules to construct more hierarchical block connections (see Figure.1(a)) to replace the FC layers with less parameters .

For the part of reconstruction unit, the reconstruction block consists of nine convolution layers followed by ReLU (see Figure.1(c)).

4. IMPLEMENTATION DETAILS

The ImageNet dataset [8] and OA-Breast dataset [9] are simulated by a paraxial approximation forward model according to Eq. (2). We use 76,000 2D simulated ImageNet data for training, 2,000 2D simulated data for validation, and 2,000 2D OA-Breast simulated images for test. All these three sets are kept separate. In our experiments, all the frequency domain data have a size of 256×128 pixels, and all the USCT images have a size of 256×256 pixels, all images are scaled to (0, 1).

5. RESULTS

We compare our method with two conventional iterative methods, L-BFGS [8] and Newton-CG [2], and two deep learning methods, U-Net [3] and FC-DenseNet103 [7], on the test set.

We use structural similarity measure (SSIM) and root-mean-square error (RMSE) for qualitative evaluation. Table 1 shows the quantitative evaluation results and the number of parameters in different methods. For the conventional algorithms, L-BFGS and Newton CG cannot achieve acceptable reconstruction results because of noisy data, and they are much slower than deep learning methods because of the expensive forward model. Compared with our proposed method, U-net and FC-DenseNet103 cannot reconstruct images well, and the reconstruction speed is slower. Our method achieves acceptable results with 0.995 SSIM and 0.027 RMSE. The total parameters in our method is 127 million, while Automap is 356 million. For Automap, it cannot reconstruct large images due to the limited GPU memory, whereas our method reduces 2/3 parameters making possible for large image reconstruction.

Table1: Quantitative evaluation of results in terms of SSIM and RMSE

Method	SoS SSIM/RMSE	Attenuation SSIM/RMSE
1) L-BFGS	0.51682/0.17413	0.04236/0.49093
2) Newton CG	0.55231/0.22754	0.04739/0.50236
3) U-Net	0.88541/0.05047	0.82313/0.05119
4) FC-DenseNet103	0.85639/0.05512	0.78951/0.06325
6) Ours	**0.99512/0.02650**	**0.99419/0.02737**

Figure 2 shows the visual comparison reconstruction results. It is obvious from the Figure 2 that our method has reached the best performance in terms of SoS and attenuation reconstruction. In addition, Figure 3 shows more visual reconstruction results of our method. In order to show more detail information, Figure 4 shows SoS and attenuation reconstruction results with soft tissue phantoms. The green, yellow, dark blue, light blue, amber color corresponds to water, skin, fat, gland and tumor, respectively.

6. CONCLUSIONS

In this paper, we propose a memory-efficient neural network architecture for USCT reconstruction, which generates high quality images. It reconstructs images from frequency domain to image domain with efficient GPU memory usage and fewer parameters. Experimental evaluations demonstrate that while state-of-the-art methods have limited capability of reconstructing USCT image from frequency domain data, our proposed method can effectively reconstruct details for USCT images, which could lead to faster diagnoses and treatment decisions for breast cancer patients.

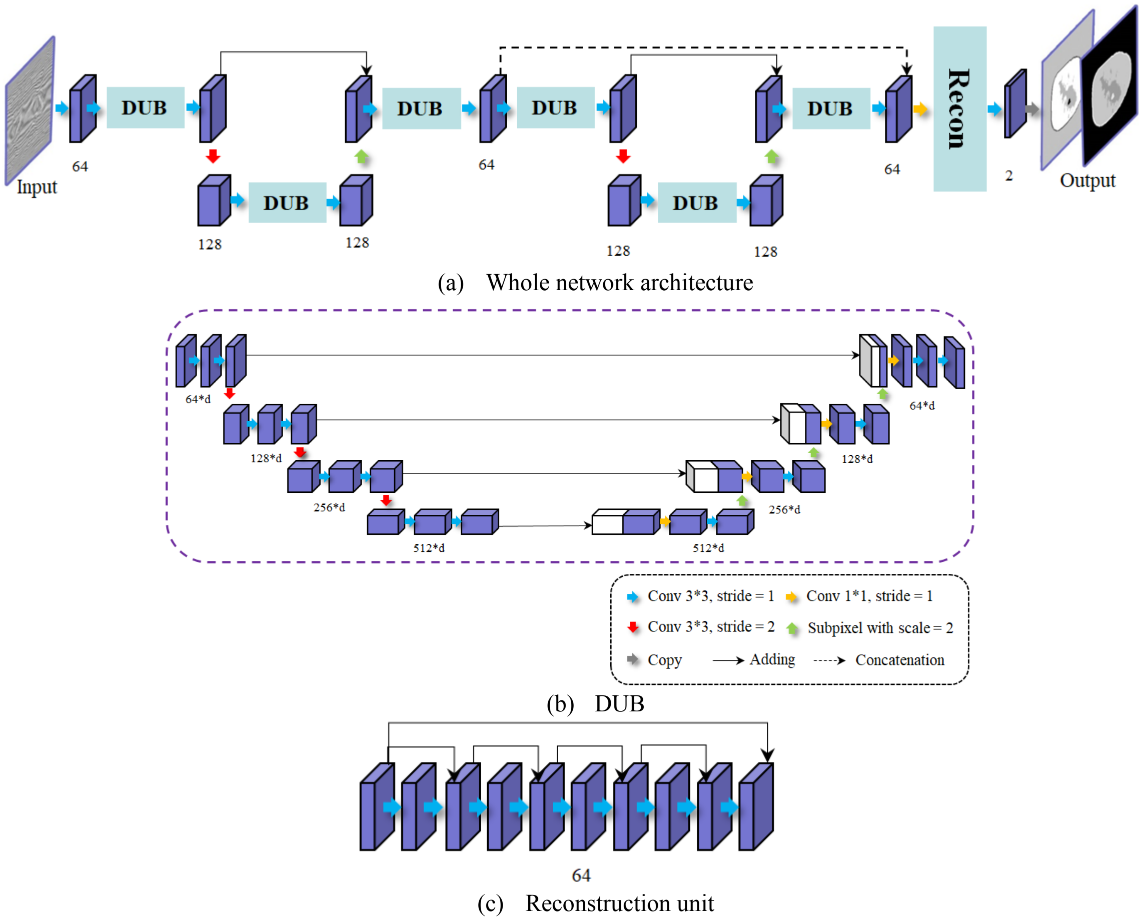

Fig.1. The domain transformation architecture for USCT reconstruction.

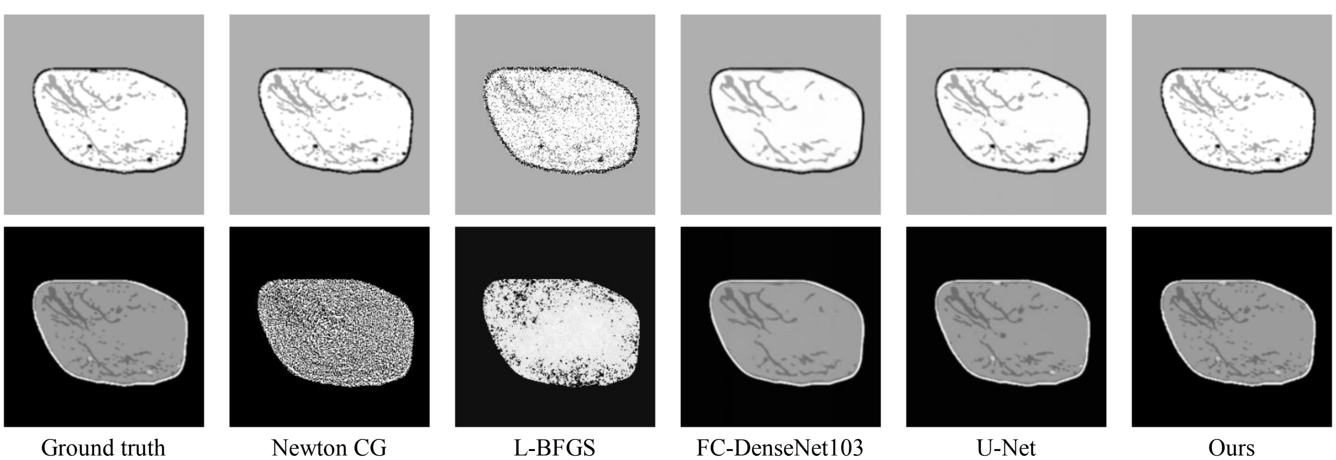

Fig.2. Visual comparisons on USCT for different algorithm. The top shows the SoS reconstruction results, while the bottom shows the attenuation results.

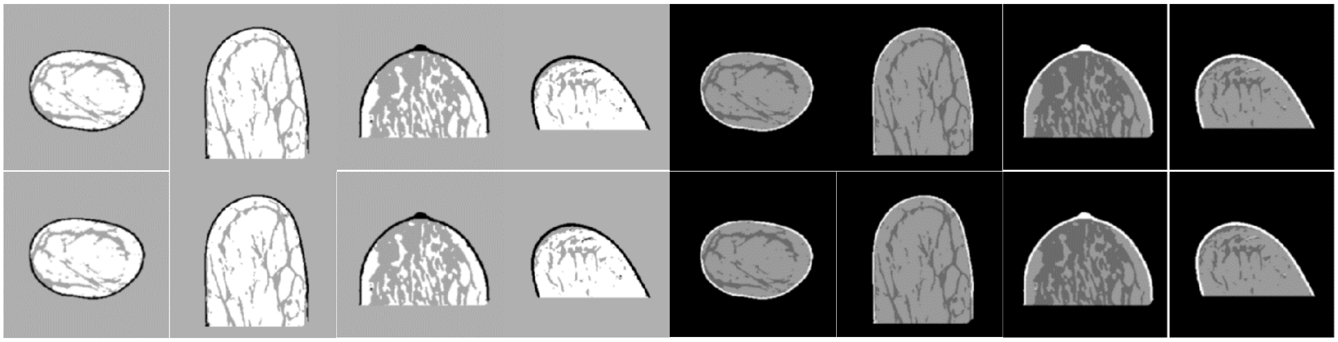

Fig.3. Visual results. The top results are the ground truth while the bottom results are the reconstructioned SoS.

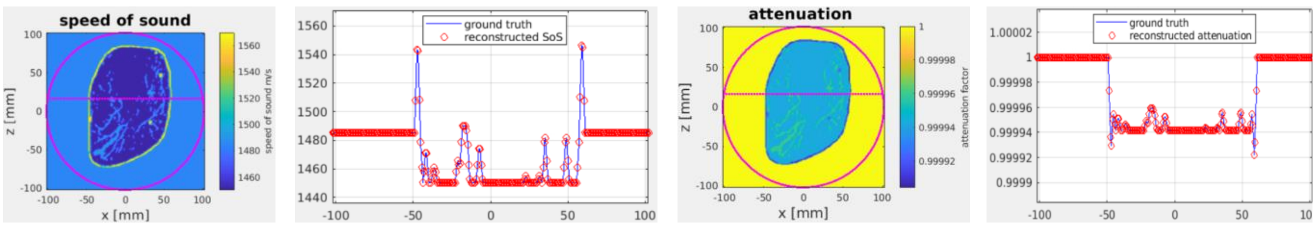

Fig.4 . Left two images: SoS reconstruction result and its 1D profiles focusing on the pixels at the pink dotted line where the reconstructed profiles are given in red circles and their simulated reference in blue. **Right two images:** Attenuation reconstruction result and its 1D profiles.

7. ACKNOWLEDGMENT

This work was supported by the Deutsche Forschungsgemeinschaft (DFG) under grants no. HE 3011/37-1 and HO 5565/2-1 and "the Fundamental Research Funds for the Central Universities" from Donghua University under grants no. 20D111205, and "the Young Teacher Research Startup Fund" from Donghua University under grants no.112-07-0053079.

8. REFERENCES

[1] H. Gemmeke, et al., "Wave equation based transmission tomography," in *Proc. IEEE Int. Ultrason. Symp. (IUS)*, pp. <1-4>, 2016.

[2] H. Wang, et al., "USCT image reconstruction: acceleration using Gauss-Newton preconditioned conjugate gradient," in *Proc. of International Workshop on Medical Ultrasound Tomography*, KIT Scientific Publishing, Germany, 2017.

[3] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. (MICCAI)*, pp. <234–241>, 2015.

[4] B. Zhu, et al., "Image reconstruction by domain-transform manifold learning." *Nature*, vol. <555>, no. <7697>, pp. < 487-492>, 2018.

[5] S.S. Girija, "TensorFlow: Large-scale machine learning on heterogeneous distributed systems.", *Software available from tensorflow. Org,* vol. <39>, no. <9>, 2016.

[6] S. Yu, B. Park, and J. Jeong, "Deep iterative down-up CNN for image denoising.", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2019.

[7] J. Simon, et al., "The one hundred layers tiramisu: Fully convolutional densenets for semantic segmentation.", *Proceedings of the IEEE conference on computer vision and pattern recognition workshops,* pp. < 487-492>, 2017.

[8] H. Wang, et al., "Accelerating image reconstruction in ultrasound transmission tomography using L-BFGS algorithm.", *Medical Imaging 2019: Ultrasonic Imaging and Tomography*, vol. < 10955>, 2019.

[9] J. Deng, et al., "Imagenet: A large-scale hierarchical image database.", *IEEE conference on computer vision and pattern recognition*, IEEE, pp. < 248-255>, 2009.

[10] L. Yang, et al., "Generation of anatomically realistic numerical phantoms for photoacoustic and ultrasonic breast imaging.", *Journal of Biomedical Optics*, vol. <22>, no. <4>, pp. <041015>, 2018.

9. COMPLIANCE WITH ETHICAL STANDARDS

This research study was conducted retrospectively using human subject data made available in open access by http://opendatacommons.org/licenses/odbl/1.0/ and http://image-net.org/. Ethical approval was not required as confirmed by the license attached with the open access data.