Program Digest, 2019 IEEE International Ultrasonics Symposium (IUS)
Glasgow, Scotland, October 6-9, 2019

A Fully Convolutional Neural Network for Rapid Displacement Estimation in ARFI Imaging

Derek Y. Chan, D. Cody Morris, Mark L. Palmeri, and Kathryn R. Nightingale
Department of Biomedical Engineering, Duke University, Durham, NC, USA
Email: derek.chan@duke.edu

***Abstract*—Ultrasound elasticity imaging in soft tissue with acoustic radiation force requires extracting displacement information, typically on the order of several microns, from raw data. In this work, we implement a fully convolutional neural network for ultrasound displacement estimation. We present a novel method for generating ultrasound training data, in which virtual displacement volumes are created with a combination of randomly-seeded ellipsoids. Network performance was tested on the virtual displacement volumes as well as an experimental phantom dataset and human *in vivo* prostate data. In simulated and phantom data, the proposed neural network accurately reconstructed the ARFI displacements, performing similarly to a conventional phase-shift displacement estimation algorithm. Application of the trained network to *in vivo* prostate data enabled the visualization of the prostatic urethra and peripheral zone.**

I. Introduction

Acoustic radiation force impulse (ARFI) imaging uses focused ultrasound to generate displacement magnitudes of several microns within tissue [1]. These displacements within the region of excitation are subsequently tracked with ultrasound, with stiffer tissues expected to have lower displacements. The efficacy of ARFI imaging has been demonstrated in several clinical applications, including cardiac, breast, and prostate imaging [2]–[4].

In ARFI imaging, 1-D in-phase and quadrature (I/Q) data are obtained before and after the induced displacements, and phase-shift autocorrelation techniques are commonly used to extract the displacement magnitudes [5]. These techniques compute the phase difference between two timesteps of interest to estimate the displacements. Repeated transmits across the field of view are used to build up a 2-D displacement image.

Machine learning methods have recently gained traction in many fields of imaging and image processing. To determine whether a neural network could learn the mapping function needed to extract displacement information from pairs of I/Q data before and after ARFI displacements, this study explored the feasibility of using a deep convolutional neural network (CNN) as an alternative to phase-shift estimators for estimating ARFI micron-level displacements.

Recent work in the literature has explored the use of deep learning for strain elastography applications. Strain elastography, which is another technique in ultrasound elastography, involves using the transducer to physically compress the tissue by several millimeters, before computing the strain to evaluate the tissue stiffness [6], [7]. Kibria and Rivaz used a neural network based on optical flow motion estimation for strain imaging [8]. Wu et al. also used a deep neural network to estimate strain for strain elastography applications [9].

The displacement estimation challenges encountered in ARFI imaging, where the displacements are orders of magnitude smaller (micron-level displacement magnitudes), are distinct from those encountered in strain imaging [5]. To our knowledge, this is the first study exploring the use of deep learning for the estimation of ARFI displacements.

Here, we introduce a novel method for generating training data in order to have a sufficiently large dataset to train the neural network. The performance of the network is evaluated in simulated, phantom, and *in vivo* human prostate data and compared with a conventional phase-shift estimator.

II. Methods

A. Training Data and Preprocessing

Simulated ultrasound data with corresponding ground-truth displacement labels were used to train the neural network. The training dataset is limited to simulated data, and not experimental ARFI acquisitions, since the true underlying displacements are known only in a simulation setting.

First, synthetic displacement fields were generated by creating 3-D volumes consisting of 150 summed ellipsoids with random size, orientation, location, and amplitude inside the volume. A 3-D Gaussian low pass filter with a standard deviation of 0.15 mm in each dimension was applied in MATLAB to prevent unrealistically sudden changes in displacement within the volume. Figure 1 shows an example of a displacement field generated from a single random seed.

Field II was used to place subresolution scatterers in a field [10], [11]. These scatterers were displaced by a given magnitude based on the 3-D displacement volumes that were generated as described previously. The software was then used to model the Siemens 12L4 ultrasound transducer (Siemens Healthcare, Mountain View, CA) to track the scatterers before and after the displacements were applied. In order to approximate the data acquired by an ultrasound scanner, the resulting raw simulated radiofrequency data were demodulated to in-phase and quadrature (I/Q) baseband frequency, and down-sampled to 5 MHz sampling frequency. The down-sampled data were then re-upsampled to 25 MHz for more precise localization of the displacements, which is a step that is typically performed prior to conventional displacement estimation [5].

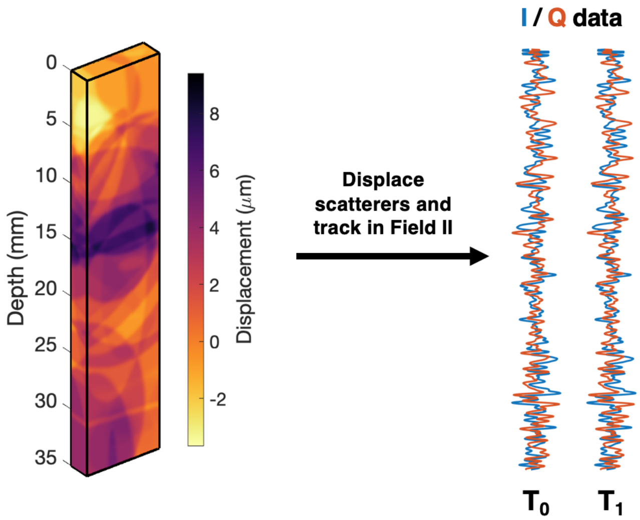

Fig. 1: Example of a synthetic 3-D displacement field generated by summing ellipsoids of random size, orientation, location, and amplitude. These displacement fields were used to displace scatterers and simulate ultrasonic tracking in Field II to produce in-phase and quadrature (I/Q) data.

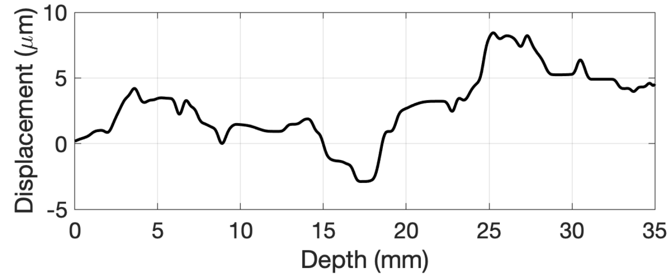

Fig. 2: Example of a ground-truth displacement label, generated by taking the center line through depth in the corresponding 3-D displacement field.

The ground truth displacement values were extracted from the center line of the virtual displacement volume, at lateral position 0 and elevation position 0. Figure 2 shows an example of the ground-truth displacements for one volume. Note that the displacement magnitudes are similar to the magnitudes typically expected in ARFI imaging applications, and the applied 3-D Gaussian spatial filter generates subtle gradients in the displacements.

Because the amplitude of the simulated Field II data is somewhat arbitrary (on the order of 10^{-21} for the simulations in this study), the data were normalized such that each A-line in the simulated data was zero-mean with unit standard deviation. This preprocessing step helps to stabilize the learning process by facilitating the weight update process during training.

In total, 30,000 different synthetic displacement fields were generated and tracked in Field II. For each sample, the two sets of I/Q data (before and after the displacements) were used as the input to the network, and the ground truth displacements through depth were used as the output to be reconstructed. The neural network was trained on 26,000 sets of these data, and 2,000 sets each were designated as validation and test datasets to evaluate the performance of the network.

B. Network Architecture and Training

Figure 2 shows a diagram of the neural network architecture. The input to the network is a M$\times$1200$\times$2$\times$2 matrix, where M is the minibatch size (set to 75), 1200 is the number of depth samples, and the last two dimensions specify the time step (before or after the applied displacements) and I/Q channel, respectively. The output is a vector of length 1200, corresponding to the displacement data through depth.

The data are then input into a series of convolutional and max pooling layers, with the number of features increasing with each convolutional layer due to the increasing complexity being represented. After four sets of convolutional layers (3$\times$3 filter size) with 2$\times$1 max pooling, a series of transposed convolutional layers are used to build the image back up to a height of 1200 samples. A final convolutional layer is used to collapse the last two dimensions to produce a 1-D output vector.

The L1 loss, or mean absolute error, was used to train the network and evaluate performance. The estimated displacements were subtracted element-wise from the ground truth displacements and the mean absolute difference across the entire minibatch was calculated. A minibatch size of 75 was used for training, with the ADAM (adaptive moment estimation) optimization algorithm and an initial learning rate of 0.001.

The network weights were initialized using the approach described by He et al., in which the variance of nodes in a layer is $2.0/n$ where n is the number of units in the previous layer [12]. This initialization was derived specifically for ReLU activation functions, which are used in this architecture, and prevents an exploding variance value as the number of inputs grows.

The fully convolutional architecture, without any fully connected layers, reduces the number of parameters in the model and is appropriate for this task since the displacement information is encoded locally (i.e., a fully connected layer would not be needed to connect distant regions of the image together).

To evaluate the displacement estimation performance of the neural network, the I/Q data were also processed using Loupas's algorithm, a conventional phase-shift displacement estimator based on a 2-D autocorrelation algorithm [5], [14]. A 1.5-wavelength kernel was used.

C. Experimental Data Acquisition

Phantom and *in vivo* prostate data were acquired using a Siemens 12L4 linear side-fire transducer on a Siemens SC2000 scanner. For an extended pushing depth of field, three focal

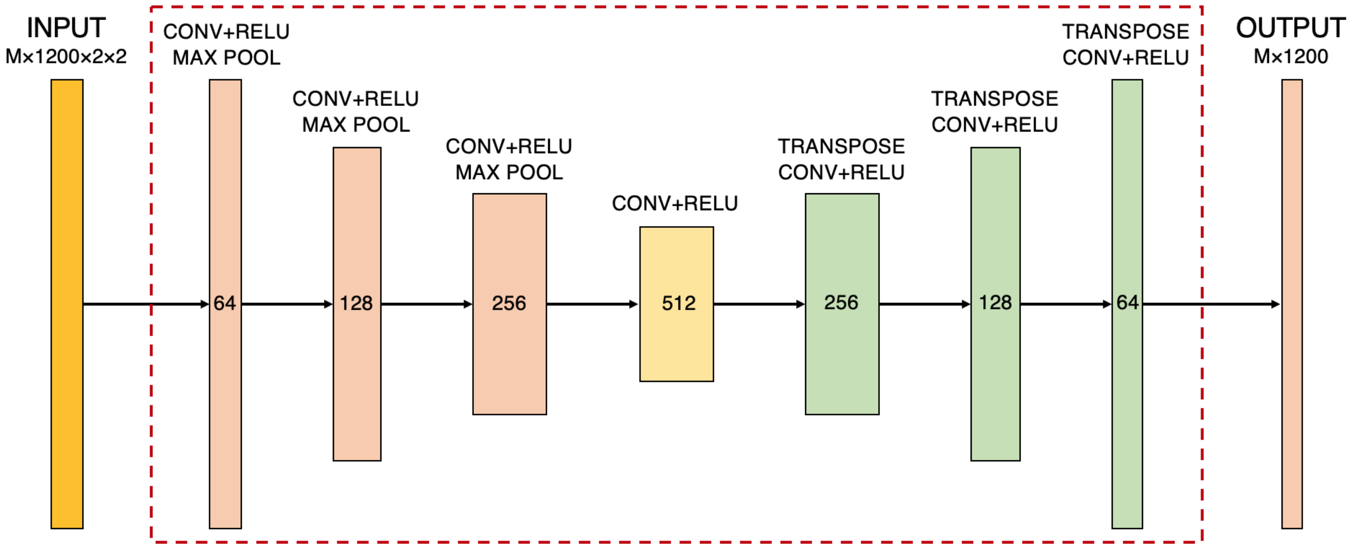

Fig. 3: Diagram of the neural network architecture. The input size is $M \times 1200 \times 2 \times 2$, where M is the minibatch size, 1200 is the number of depth samples, and the last two dimensions specify the time step and I/Q data channel. A series of convolutional and max pooling layers are used, followed by a series of transposed convolutional layers to build the image size back up to 1200 samples. A final convolutional layer collapses the last two dimensions to produce a 1-D displacement output vector.

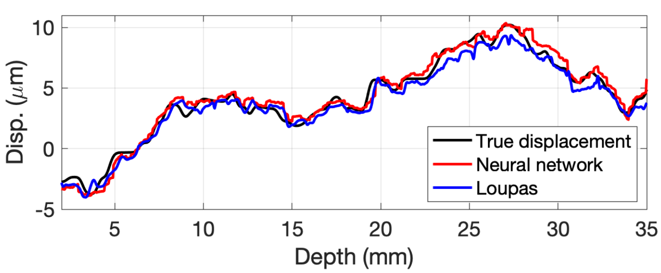

Fig. 4: Displacement estimation results for simulated data generated from synthetic 3-D displacement field.

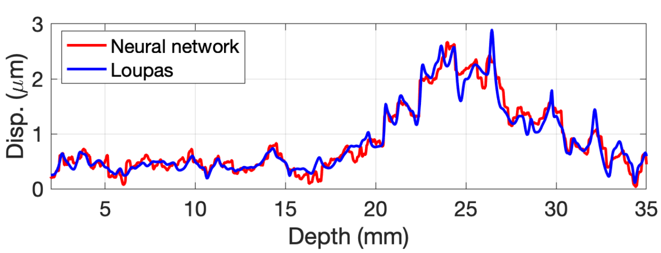

Fig. 5: Displacement estimation results for experimental data acquired in a phantom.

depths were used for each radiation force excitation: 30 mm, 22.5 mm, and 15 mm [13]. The track transmit beam was focused at 60 mm in an F/3 configuration with dynamic receive focusing, and eighty-two push beams were laterally transmitted across the aperture to produce a 2-D imaging plane. The phantom was a custom CIRS elastic phantom (Norfolk, VA) containing a stiff spherical inclusion with a

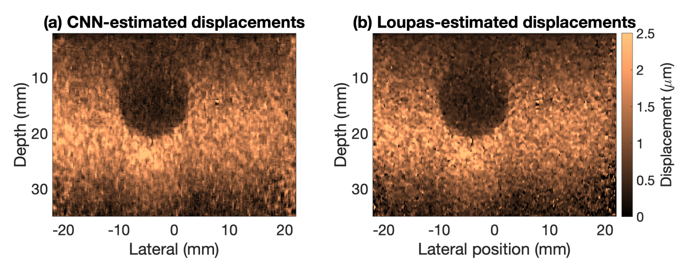

Fig. 6: Displacement estimation results for experimental data acquired in a phantom (RMS difference = 0.40 μm)

diameter of 10 mm.

In an institutional review board-approved study, prostate data were obtained in men before they underwent a radical prostatectomy procedure [4]. The same ARFI sequence described above was used, and for each subject, a 3-D prostate data volume was populated by rotating the side-fire transducer in 1 degree increments in elevation using a mechanical rotation stage with an optical encoder to track the trajectory of the transducer.

III. Results and Discussion

Figure 4 shows displacement estimation results in a simulated dataset (i.e., tracked data from one of the random 3-D displacement fields in the test dataset). The ground-truth displacements based on the central line of the displacement field are shown in black, while the CNN-estimated dipslacements and Loupas-estimated displacements are shown in red and blue, respectively. Both the neural network and Loupas's algorithm were able to reconstruct the displacements

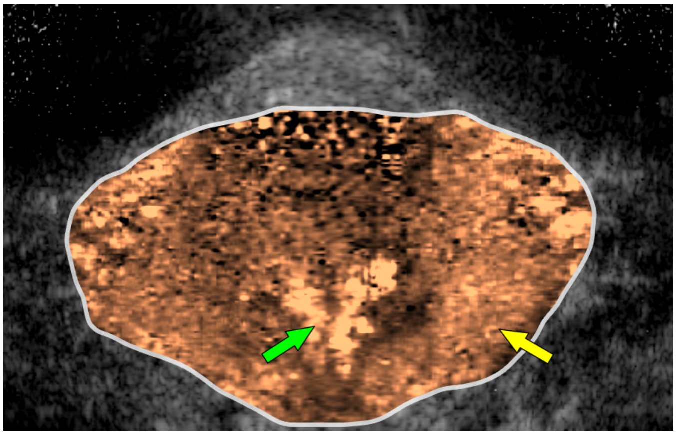

Fig. 7: Coronal *in vivo* prostate ARFI image obtained using the neural network. The green arrow indicates the prostatic urethra, while the yellow arrow points to the peripheral zone of the prostate.

from the simulated I/Q data. The root mean squared error (RMSE) between the CNN displacements and ground truth was 0.54 μm, while the RMSE was 0.61 μm for the Loupas displacements.

Deviations from the ground-truth displacements in Figure 4 are likely due to an averaging effect of the imaging point spread function (”shearing”) [15]. In other words, scatterer displacements are inhomogeneous within the ultrasound track beam and the different displacements are averaged together in the raw ultrasound data, which makes it extremely challenging to determine the true displacement of the scatterers along the central axis.

The neural network generalized to experimental phantom data, as shown in Figure 5. Again, the CNN-estimated displacements and Loupas-estimated displacements are shown in red and blue, respectively, and there was no ground-truth for comparison since this was an experimental acquisition. The root mean squared difference between the CNN displacements and Loupas-estimated was 0.24 μm, indicating very concordant estimates. Figure 6 shows images from the same experimental phantom data, where the stiff spherical inclusion is visualized as a circular region of low displacement. Again, the convolutional neural network and Loupas produce consistent displacement images (RMS difference = 0.40 μm).

Figure 7 shows results of applying the convolutional neural network displacement estimator to a human prostate dataset. The image shows a scan-converted coronal plane from the 3-D ARFI volume. Prostate anatomy is clearly visualized in this ARFI image: the green arrow indicates the prostatic urethra, while the yellow arrow points to the peripheral zone of the prostate.

IV. Conclusions

In this study, a fully convolutional neural network was trained to extract ARFI small displacements from ultrasound data, using a novel method for generating synthetic 3-D displacement volumes that were tracked in simulations to produce the training dataset. In simulated data, the network accurately reconstructed the ground-truth displacements. The trained network generalized to experimentally-acquired phantom data, enabling the visualization of a stiff spherical inclusion contained within an elastic phantom. Using the neural network in human *in vivo* prostate data, the peripheral zone and urethra were well-visualized.

Acknowledgments

This work was supported by NIH grant R01-CA142824 and DOD PCRP grant W81XWH-16-1-0653. The authors thank Siemens Medical Solutions USA, Ultrasound Division for in-kind and technical support and Ned Danieley for computer system administration.

References

[1] J. Doherty, G. E. Trahey, K. . Nightingale, and M. L. Palmeri, “Acoustic radiation force elasticity imaging in diagnostic ultrasound,” IEEE Trans. Ultrason. Ferroelectr. Freq. Control, vol. 60, no. 4, pp. 685–701, Apr. 2013.

[2] P. Hollender et al., “Intracardiac acoustic radiation force impulse (ARFI) and shear wave imaging in pigs with focal infarctions,” IEEE Trans. Ultrason. Ferroelectr. Freq. Control, vol. 60, no. 8, pp. 1669–1682, Aug. 2013.

[3] S.-D. Wang et al., “Differential diagnostic performance of acoustic radiation force impulse imaging in small (20 mm) breast cancers: is it valuable?” Sci. Rep., vol. 7, no. 8650, 2017.

[4] M. L. Palmeri et al., “Identifying clinically significant prostate cancers using 3-D in vivo acoustic radiation force impulse imaging with whole-mount histology validation,” Ultrasound Med. Biol., vol. 42, no. 6, pp. 1251–1262, June 2016.

[5] G. F. Pinton, J. J. Dahl, and G. E. Trahey, “Rapid tracking of small displacements with ultrasound,” IEEE Trans. Ultrason. Ferroelectr. Freq. Control, vol. 53, no. 6, pp. 1103–1117, June 2006.

[6] P. N. Wells and H. D. Liang, “Medical ultrasound: imaging of soft tissue strain and elasticity,” J. R. Soc. Interface, vol. 8, no. 64, pp. 1521–1549, Nov. 2011.

[7] R. M. S. Sigrist et al., “Ultrasound elastography: review of techniques and clinical applications,” Theranostics, vol. 7, no. 5, pp. 1303–1329, Mar. 2007.

[8] M. G. Kibria and H. Rivaz, “Global ultrasound elastography using convolutional neural network,” 21st Med. Image. Comput. Comput. Assist. Interv., Granada, Spain, Sept. 2018.

[9] S. Wu et al., “Direct reconstruction of ultrasound elastography using an end-to-end network,” 21st Med. Image. Comput. Comput. Assist. Interv., Granada, Spain, Sept. 2018.

[10] J. A. Jensen, “Field: a program for simulating ultrasound systems,” Med. Bio. Eng. Comput., vol. 34, no. 1, pp. 351–353, 1996.

[11] J. A. Jensen and N. B. Svendsen, “Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers”, IEEE Trans. Ultrason. Ferroelectr. Freq. Control, vol. 39, no. 2, pp. 262–267, Mar. 1992.

[12] K. He, X. Zhang, S. Ren, and J. Sun, “Delving deep into rectifiers: surpassing human-level performance on ImageNet classification,” Proc. ICCV, Santiago, Chile, Dec. 2015.

[13] J. Bercoff, M. Tanter, and M. Fink, “Supersonic shear imaging: a new technique for soft tissue elasticity mapping,” IEEE Trans. Ultrason. Ferroelectr. Freq. Control, vol. 51, no. 4, pp. 396–409, Apr. 2004.

[14] 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, July 1995.

[15] M. L. Rosenzweig, S. A. McAleavey, Gregg E. Trahey, and K. R. Nightingale, “Ultrasonic tracking of acoustic radiation force-induced displacements in homogeneous media,” IEEE Trans. Ultrason. Ferroelectr. Freq. Control, vol. 53, no. 7, pp. 1300–1313, Jul. 2006.